Finite-Length Scaling Laws for Iterative Coding Systems

Finite-Length Scaling Laws for Iterative Coding Systems

Jérémie Ezri

EPFL - Ecole Polytechnique Fédérale de Lausanne

Jérémie Ezri
21/8/1981 – 20/4/2010

In late 2007, Jérémie Ezri started to write down the first notes of what he hoped would become his thesis. Soon thereafter, in mid January 2008, he was diagnosed with cancer. Over the following two years, through the ups and downs of surgery and chemo therapy, he continued to document his research. Jérémie died on April 20, 2010. This booklet contains his thesis. It is essentially in the form Jérémy had left it, a few months away from his defense.

Besides the results found in these notes, his scientific contributions also include material which was published in the following papers:

J. Ezri and M. Gastpar, "On the performance of independently designed LDPC codes for the relay channel", ISIT, 2006, USA, 2006, pp. 977–981.

J. Ezri, A. Montanari, and R. Urbanke, "Finite-Length Scaling for Gallager A", 44th Allerton Conference on Communication, Control, and Computing, 2006, USA.

J. Ezri, A. Montanari, and R. Urbanke, "A generalization of the finite-length scaling approach beyond the BEC", ISIT, 2007, France, pp. 1011–1015.

J. Ezri, A. Montanari, S. Oh, and R. Urbanke, "Computing the threshold shift for general channels", ISIT, 2008, USA, pp. 1448–1452.

J. Ezri, A. Montanari, S. Oh, and R. Urbanke, "The slope scaling parameter for general channels, decoder, and ensembles", ISIT, 2008, USA, pp. 1443–1447.

For his thesis, Jérémie worked on an ambitious project with the same enthusiasm and passion that he had for music and for his hobby of creating movies. When he started, there was little more than a general vision – how to make the optimization of iterative codes accessible to a general audience. The idea, inspired from statistical physics, was to derive scaling laws to describe the difficult-to-evaluate finite-length performance of a code using quantities that are computable in the limit of infinitely large block lengths. At the time, only the simplest case had been solved, namely the transmission over the so called binary erasure channel. It was far from clear whether the necessary extensions could be carried out.

Over time, working his way methodically from simple extensions to the general case, he showed that the general approach was not only possible, but quite promising. By developing the necessary theory and implementing the associated algorithms, he was able to accurately predict the performance of modern codes. A quick look at the daunting mathematical expressions in this thesis should make it clear what a difficult task this was and how impressive it is that he succeeded.

Working with Jérémie never felt like work. Not even the pressure and the stress preceding an important deadline would make him lose his mischievous side. His impromptu imitations were legendary. Whether summer intern or visiting Shannon award winner – no one was spared. He was also one of the founding members of the IPG movie production team.

We have lost a great researcher, the most likely IPG member to make it to Star Academy, our SuperStar, and a great friend – all at once.

Ruediger Urbanke
(for all of IPG)

Abstract

In the past 20 years, codes on graphs have found their way from academic research into everyday communication systems. Most modern communications standards use codes on graphs as their main means of establishing an efficient and reliable link.

The literature on codes on graphs is substantial, with most of it focusing on code constructions and their analysis when the block length tends to infinity. This has lead to a large number of known constructions which all perform well in the asymptotic limit. But how do these codes perform for "practical" lengths? This is a challenging question which has considerable practical relevance.

Scaling laws are perhaps the most promising avenue for answering this question. For codes on graphs the idea of scaling laws was previously explored for the binary erasure channel and it was shown in this case how to compute the scaling parameters and how to accomplish a finite-length code optimization.

We show that a similar approach can be carried through for general binary memoryless symmetric channels. More precisely, we investigate the form of the scaling laws as well as means of computing the scaling parameters for various iterative decoding algorithms and irregular low- density parity-check codes. We first show how to compute the message variance for a fixed number of iterations for irregular low-density parity-check ensembles. From these calculations the basic scaling parameter α can be deduced. By means of examples we demonstrate that the predicted performance is a very good indicator for the actual measured data. This opens up the way of using scaling laws in a constructive fashion, as part of a finite-length optimization routine.

Our aim is to provide an easily usable finite-length optimization tool that is applicable to the wide variety of channels, blocklengths, error probability requirements, and decoders that one encounters in practical systems. The tool is aimed at non-experts in the field, who need to quickly find code designs that are comparable with the best known codes available today but do not have the luxury of spending months in doing so.

Résumé

En 20 ans, les codes sur graphes sont passés de la recherche académique à des applications concrètes dans les systèmes de communication utilisés quotidiennement. En particulier, les standards de communication moderne utilisent les codes sur graphes comme un moyen privilégié d'établir un canal de communication sûr et de qualité.

Il existe de nombreux travaux relatifs aux codes sur graphes. Ces travaux se concentrent en grande partie sur la construction de ces codes ainsi que sur leur analyse lorsque la taille des blocs tend vers l'infini. Cela conduit à de multiples constructions de codes avec de bonnes performances asymptotiques. Mais quelle est la performance de ces codes pour les tailles de blocs utilisées en pratique? Il s'agit d'une question primordiale d'une importance considérable.

Les lois d'échelle sont probablement la piste la plus prometteuse pour essayer de répondre à cette question. Dans le passé, cette approche a été appliquée au cas du canal binaire à effacement pour lequel il a été montré comment calculer les paramètres d'échelle et optimiser les codes de tailles finies.

Nous montrons qu'une approche similaire peut être utilisée pour l'ensemble des canaux binaires, symétriques, sans mémoire. Plus précisément, nous étudions la forme des lois d'échelle et les facons de calculer les paramètres d'échelle pour différents types d'algorithmes itératifs de décodage et codes LDPC irréguliers. Dans un premier temps, nous montrons comment calculer la variance des messages pour un nombre fixé d'itérations, pour les ensembles LDPC irréguliers. De ces calculs nous pouvons également déduire le paramètre d'échelle de base, α. Nous utilisons ensuite des exemples pour démontrer que les performances prédites sont très proches des données mesurées. Ceci ouvre la porte pour l'utilisation des lois d'échelle dans le contexte de l'optimisation des codes de taille finie.

Notre but est d'offrir un outil d'optimisation des codes de taille finie facile a utiliser et applicable a une grande variété de scénarios en terme de canaux, taille de blocs, seuil de probabilité d'erreur et décodeur. Cet outil vise des utilisateurs non-experts, désireux de construire rapidement des codes aux performances comparables celles des meilleurs codes disponibles sur le marché.

Contents

Introduction 1

1.1 History of Scaling Laws for LDPC Codes

Assume that you want to optimize a sparse graph code under iterative decoding for a fixed length. In general, it is not optimal to choose the code with the best *asymptotic* (in the blocklength) performance since the convergence speed of different ensembles to this asymptotic limit varies considerably and this difference can therefore not be ignored. We are hence looking for good finite-length approximations of the performance. If the approach is to be of any practical value it must be quite flexible, amenable to the analysis of the large collection of ensembles as well as to the variety of message-passing decoders (quantizations of BP) that have been proposed to date.

One possible approach is to use *scaling laws*, which have a long and successful history in statistical physics. We refer the reader to the books by Fisher [1] and Privman [2]. The idea of scaling laws was introduced into the coding theory literature by Montanari [3]. By means of a specific example (regular ensemble, binary symmetric channel (BSC) with parameter ϵ) he showed that if one plots the block error probability as a function of the "scaled variable" $z = \sqrt{n}(\epsilon^{\mathrm{BP}} - \epsilon)$ (where ϵ^{BP} is the density evolution threshold under belief propagation (BP)), then the curves corresponding to increasing blocklengths quickly converge to a single "mother curve", called the *scaling function.* The results of [3] are repeated in Figure 1.1.

This suggested that if one were able to analytically determine the scaling function as well as the scaling parameters for a given system (degree distribution, channel and decoder) then it would be possible to efficiently and accurately predict the finite-length performance of iterative coding systems.

The first analytic result was derived by Amraoui, Montanari, Richardson, and Urbanke [4, 5]. They showed that, for transmission over the binary erasure

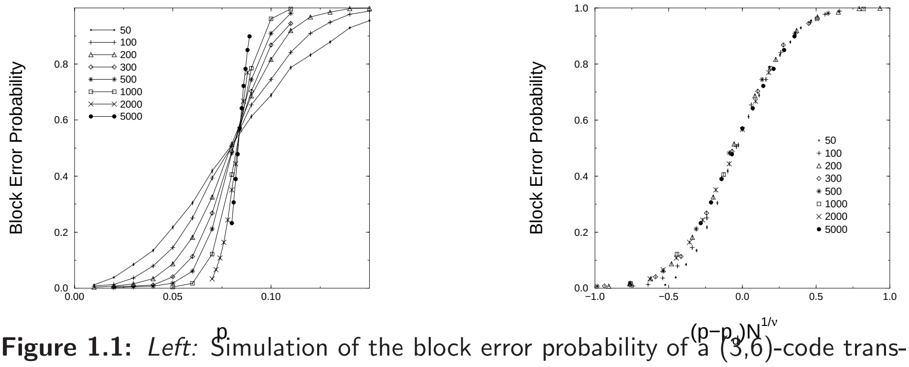

Figure 1.1: *Left:* Simulation of the block error probability of a (3,6)-code transmitted over a BSC and decoded with a belief propagation decoder. The corresponding threshold is $\epsilon^{\text{BP}} = 0.084$. *Right:* Rescaled block error probability with respect to $z = \sqrt{n}(\epsilon - \epsilon^{\text{BP}})$. Note that the simulation points all cluster around a single *mother* curve. Both figures are taken from [3].

channel (BEC) with parameter ϵ, the block error probability behaves like

$$P_B(z) = Q\Big(\frac{z}{\alpha}\Big)(1 + o_n(1)),$$

where $z = \sqrt{n}(\epsilon^{\text{BP}} - \epsilon)$ and where ϵ^{BP} is the threshold of the ensemble under BP decoding. They further conjectured that, more accurately,

$$P_B(z) = Q\Big(\frac{z}{\alpha}\Big)(1 + O(n^{-\frac{1}{3}})), \tag{1.1}$$

where $z = \sqrt{n}(\epsilon^{\text{BP}} - \beta n^{-\frac{2}{3}} - \epsilon)$, and where the term $\beta n^{-\frac{2}{3}}$ represents a *finite-length shift* of the threshold. More precisely, $\epsilon^{\text{BP}}(n) - \epsilon^{\text{BP}}(\infty) = \beta n^{-\frac{2}{3}} + o_n(1)$, where $\epsilon^{\text{BP}}(n)$ is the finite-length threshold, i.e., the channel parameter where the average block error probability of the ensemble of length n takes on the value one-half.

Analytically the scaling law promises the convergence of $P_B(z)$ for a fixed z and n tending to infinity. In practice, the scaling law provides accurate predictions already for moderate blocklengths and also away from the threshold; see e.g., the left-hand side graph in Figure 3.2. The analysis put forward in [4, 5] was based on the so called *peeling decoder* by Luby, Mitzenmacher, Shokrollahi, Spielman, and Steman, see [6]. This decoder is equivalent to the standard BP decoder but represents the decoding process as a sequence of discrete steps, where each step corresponds to determining a previously erased bit from its known neighbors. It was shown in [4, 5] how to determine the scaling parameter α from the solution of a system of differential equations, which were dubbed covariance evolution. These equations were then solved numerically in order to evaluate α.

An alternative way to determine α for the BEC case was proposed in [7]. Recall that density evolution computes the *average* number of erased messages as a function of the iteration number. It was shown in [7] that α can be computed by determining the *variance* of the number of variable-to-check messages. This computation was accomplished in [8] and an explicit value for α as a function of the degree distribution pair (λ, ρ) was given. Further, explicit expressions for β were derived. Given these explicit expressions of the scaling parameters, it is then easy to accomplish a finite-length optimization.

1.2 Thesis Outline

All the above developments were restricted to the BEC. In Chapter 2 we take the first step towards extending the scaling law to more general channels. More precisely, we consider transmission over the BSC and decoding via the Gallager algorithm A. This algorithm has the property that for most ensembles the threshold is given by a fixed point at the beginning of the decoding process. This somewhat simplifies the determination of the scaling law and the scaling parameter. During our investigation we will see an interesting new feature of general message-passing decoders – if we consider density evolution on an infinite tree above threshold, then even though the message densities converge, the messages themselves do not. This makes the analysis more interesting but also much more challenging. A detailed analysis of this general phenomenon, which we termed “flipping,” is contained in Chapter 4.

In Chapter 3 we then consider a general setup. More precisely, we define a broad class of quantized message-passing algorithms as well as an “EXIT-like” curve which characterizes the decoding performance. Under some mild regularity condition on this EXIT-like curve, we then show that it is possible to extend the alternative derivation of the scaling law for the BEC which was introduced by Ezri, Montanari, and Urbanke in [9].

2 Scaling Law for Gallager A

2.1 Introduction

As a first step in generalizing scaling laws to more general channels we consider transmission over the BSC using the simplest message-passing decoder, namely, Gallager A. To simplify things further, we only consider ensembles whose threshold is given by a fixed point at the start of the decoding process.

2.2 Gallager A Algorithm: A Short Review

Gallager introduced in his thesis [10] what is now called Gallager's decoding algorithm A. Messages are from the set $\{0, 1\}$ and represent the current estimate of the decoder of a particular bit. At check nodes, the message-passing rules call for the computation of the `XOR` sum of the incoming messages. At a variable node the outgoing message equals the originally received message, except if all *other* incoming messages agree, in which case this common value is sent.

The density evolution equations corresponding to this decoder were already written down by Gallager. By convention, one round of message passing starts with the processing at the check nodes, is followed by sending the message to the variable nodes, then comes the processing at the variable nodes, and, it concludes, by sending the messages from the variable nodes to the check nodes. Let y_ℓ/x_ℓ denote the error probability of the check-to-variable/variable-to-check messages in iteration ℓ. Define

$$\begin{aligned} g(x) =& \frac{1-(1-2x)^{\mathtt{r}-1}}{2}, \\ f(\epsilon, x) =& \epsilon(1-(1-g(x))^{\mathtt{l}-1}) + (1-\epsilon)g(x)^{\mathtt{l}-1}. \end{aligned}$$

We have $x_0 = \epsilon$. Then for $\ell \geq 1$, $y_\ell = g(x_{\ell-1})$ and $x_\ell = f(\epsilon, y_\ell)$. Various analytic properties of these equations and methods for an exact computations of the threshold can be found in [11].

The threshold of a regular LDPC ensemble under Gallager A is determined by one of the following three conditions: (i) the stability condition which reads $\epsilon^{\text{Sta}} = \frac{1}{(\mathtt{l}-1)(\mathtt{r}-1)}$ (see [11]), (ii) a fixed point in the middle of the decoding process, or (iii) a fixed point at the beginning of the decoding process, $f(\epsilon, \epsilon) = \epsilon$. The last case is unusual in the sense that it does not happen under BP decoding. But it was shown in [11] that (i) and (ii) are the most common cases under Gallager A. Figure 2.1 demonstrates these three cases by means of three regular ensembles.

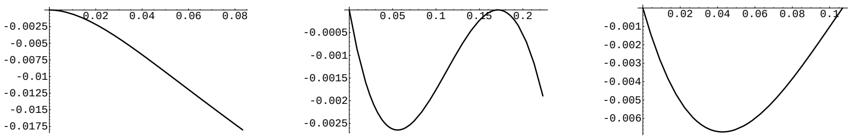

Figure 2.1: Left: the function $f(\epsilon^{\mathsf{BP}}, x)$ for the $(4, 5)$ regular code whose threshold is given by the stability condition $\epsilon^{\mathsf{BP}} = \epsilon^{\mathsf{Sta}} = \frac{1}{12}$. Middle: the function $f(\epsilon^{\mathsf{BP}}, x)$ for the $(3, 3)$ regular code whose threshold is given by a "regular" fixed point and equals $\epsilon^{\mathsf{BP}} \approx 0.2230467$. Right: the function $f(\epsilon^{\mathsf{BP}}, x)$ for the $(3, 4)$ regular code whose threshold is given by a fixed point at the start of the decoding process – $f(\epsilon^{\mathsf{BP}}, \epsilon^{\mathsf{BP}}) = \epsilon^{\mathsf{BP}}$, $\epsilon^{\mathsf{BP}} \approx 0.106924$.

One would expect that the scaling is quite different in all these three cases. Indeed, under BP decoding and the BEC the scaling for cases (i) and (iii) were investigated in [4] and found to be of fundamentally different nature. We investigate the scaling for case (iii). Somewhat surprising, the scaling law has the same form as the case (ii) under BP decoding for the BEC.

2.3 Derivation of Scaling Law for Gallager A

We restrict our attention to ensembles whose threshold is determined by a fixed point right at the beginning of the decoding process. Consider the following experiment. Pick a code and a channel realization at random. Let X_ℓ denote the number of erroneous variable-to-check messages in iteration ℓ. Since the graph and the channel realization are random, this number is also a random variable. Fix ℓ to some integer and consider the following block error probability estimator: if $X_\ell - X_0 > 0$, i.e., if the number of errors has increased after ℓ iterations then we estimate that the decoding process will fail (the actual decoder might run for many more iterations). Otherwise we assume that it will succeed (breaking ties in a fixed but arbitrary way). Table 2.1 shows how good this estimate performs for two regular ensembles and 1, 2, and 4 iterations. As we can see from this data, even after one iteration the estimate is not too bad

Ensemble	1 iter.	2 iter.	4 iter.
$(3,6)$	$85,2\%$	$90,7\%$	$96,6\%$
$(3,4)$	$82,3\%$	$85,8\%$	88.2%

Table 2.1: Performance of block error probability estimator based on observing only the first ℓ iterations for the $(3,6)$-regular and the $(3,4)$-regular ensemble. The blocklength is $n = 1000$. We see that e.g. in roughly 85% of the cases if the number of erroneous messages decreased (increased) after the first iteration then the final decoding was successful (the decoder failed).

and it improves as we increase the iteration number. Of course, if we fix the length of the code and let ℓ increase to the true number of iterations of the decoder then this estimator is exact.

Let us derive a scaling law for this estimator for a fixed ℓ. Define $Z_\ell(\epsilon) = \frac{X_\ell - X_0}{\sqrt{n\mathtt{1}}}$. Compute

$$\sigma_\ell^2(\epsilon) = \mathbb{E}[Z_\ell(\epsilon)^2] - \mathbb{E}[Z_\ell(\epsilon)]^2.$$

It is not very hard to show that if we fix ℓ then Z_ℓ converges in distribution (as a function of the length n) to a Gaussian. We have $\lim_{n\to\infty} \mathbb{E}[Z_\ell(\epsilon^{\mathrm{BP}})] = 0$ and by definition $\lim_{n\to\infty} \mathbb{E}[Z_\ell(\epsilon^{\mathrm{BP}})^2] = \sigma_\ell^2(\epsilon^{\mathrm{BP}})$. Imagine that we change ϵ by $\Delta\epsilon$, a very small quantity (our language below will reflect the case where $\Delta\epsilon > 0$). Note that $\sigma_\ell^2(\epsilon)$ is a continuous function of ϵ so that the variance remains essentially unchanged by this small change in ϵ. But because of the change of ϵ the expected value of Z_ℓ changes from zero to

$$\lim_{n\to\infty} \frac{1}{\sqrt{n\mathtt{1}}} \frac{\mathrm{d}\mathbb{E}[Z_\ell(\epsilon)]}{\mathrm{d}\epsilon} \Delta\epsilon = \frac{\mathrm{d}z_\ell(\epsilon)}{\mathrm{d}\epsilon} \Delta\epsilon,$$

where $z_\ell(\epsilon) = x_\ell - x_0$. This means that *on average* the number of errors increases after one iteration. Nevertheless, there are *instances* for which the number of errors *decreases*. More precisely, this happens with probability $Q\left(\frac{\sqrt{n\mathtt{1}}\frac{\mathrm{d}z_\ell(\epsilon)}{\mathrm{d}\epsilon}\Delta\epsilon}{\sigma_\ell(\epsilon)}\right)$, where we have made use of the fact that Z_ℓ converges in distribution to a Gaussian. This gives us the scaling law of our estimator.

Theorem 2.1 (Scaling Law for Estimator)**.** *Consider a (λ, ρ) degree distribution pair whose threshold under Gallager A is given by a fixed point at the beginning of the decoding process. Let $\mathtt{G}$ denote an element of the ensemble LDPC(n, λ, ρ). Then for $w = \sqrt{n}(\epsilon - \epsilon^{BP})$ fixed and n increasing*

$$P(Z_\ell(\mathtt{G}, \epsilon) > 0) = Q\left(\frac{w}{\alpha_\ell}\right)(1 + o(1)),$$

where

$$\alpha_\ell = \left.\frac{\sigma_\ell(\epsilon)}{\sqrt{\mathtt{1}}\frac{dz_\ell(\epsilon)}{d\epsilon}}\right|_{\epsilon=\epsilon^{BP}}.$$

Conjecture 2.1 (Scaling Law For Gallager A)**.** *Consider a (λ, ρ) degree distribution pair whose threshold under Gallager A is given by a fixed point at the beginning of the decoding process. Let* G *denote an element of the ensemble LDPC(n, λ, ρ). Then for $w = \sqrt{n}(\epsilon - \epsilon^{BP})$ fixed and n increasing*

$$P_B(w) = Q\left(\frac{w}{\alpha}\right)(1 + o(1)),$$

where

$$\alpha = \lim_{\ell \to \infty} \alpha_\ell = \lim_{\ell \to \infty} \left. \frac{\sigma_\ell(\epsilon)}{\sqrt{1}\frac{dz_\ell(\epsilon)}{d\epsilon}} \right|_{\epsilon = \epsilon^{BP}}.$$

Discussion: Note that we have phrased the scaling law regarding the error probability of the real decoder as a conjecture. The reason is simple: the scaling of the true probability of error follows by tracing the behavior if we first let the number of iterations tend to infinity and then let the block length grow. Our computations on the other hand are derived by exchanging the order of limits and we have not shown that the scaling stays unchanged under this exchange.

There are two remaining tasks. The first one is dispensed with rather quickly.

Lemma 2.1 (Computation of $\frac{\mathrm{d}z_\ell(\epsilon)}{\mathrm{d}\epsilon}$)**.**

$$\left. \frac{\mathrm{d}z_\ell(\epsilon)}{\mathrm{d}\epsilon} \right|_{\epsilon = \epsilon^{BP}} = \frac{b + (1 - a - b)a^\ell}{1 - a} - 1,$$

where $a = \frac{\partial f(\epsilon, x)}{\partial x} |_{\epsilon = x = \epsilon^{BP}}$ and $b = \frac{\partial f(\epsilon, x)}{\partial \epsilon} |_{\epsilon = x = \epsilon^{BP}}$.

Proof.

$$\begin{aligned}
&\frac{\mathrm{d}z_\ell(\epsilon)}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} \\
&= \frac{\mathrm{d}(x_\ell(\epsilon) - x_0(\epsilon))}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} = \frac{\mathrm{d}(f(\epsilon, x_{\ell-1}(\epsilon)) - \epsilon)}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} \\
&= \frac{\mathrm{d}f(\epsilon, x_{\ell-1}(\epsilon))}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} -1 \\
&= \frac{\partial f(\epsilon, x)}{\partial \epsilon} \,|_{\epsilon = x = \epsilon^{\mathrm{BP}}} + \frac{\partial f(\epsilon, x)}{\partial x} \,|_{\epsilon = x = \epsilon^{\mathrm{BP}}} \frac{\mathrm{d}x_{\ell-1}(\epsilon)}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} -1 \\
&= b + a\frac{\mathrm{d}x_{\ell-1}(\epsilon)}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} -1 \\
&= b + a\frac{\mathrm{z}_{\ell-1}(\epsilon)}{\mathrm{d}\epsilon} \,|_{\epsilon = \epsilon^{\mathrm{BP}}} +a - 1.
\end{aligned}$$

By an explicit calculation we can check that the given solution fulfills this recursion: indeed

$$\begin{aligned}
&b + a\left(\frac{b + (1 - a - b)a^{\ell-1}}{1 - a} - 1\right) + a - 1 \\
&= \frac{b(1 - a - b)a^\ell}{1 - a} - 1.
\end{aligned}$$

□

The second task is to determine σ_ℓ^2. This is much more involved. We show how to compute this quantity for *regular* ensembles in Section 2.6.

2.4 Comparison to Simulation Results

Example 2.1. *If we evaluate the recursions given in Section 2.6 we get for* $\ell = 1, \cdots, 11$, $\sigma_\ell^2(3,4) = 0.13274134, 0.51677864, 1.23228379, 2.43133296, 4.28314003, 7.04461, 11.0377, 16.7088, 24.6399, 35.6128, 50.7295$. *This corresponds to* $\alpha_\ell(3,4) = 1.22027, 1.122, 1.07487, 1.05214, 1.03644, 1.02608, 1.01828, 1.01248, 1.00792, 1.00432, 1.00211$. *We see that* $\alpha_\infty(3,4) \approx 1.0$.

In a similar manner the consecutive values of $\sigma_\ell^2(3,6)$ *are* $0.0753008, 0.393956, 1.218374, 3.1463, 7.336317, 16.1522, 34.27, 70.9914, 144.5142, 292.62843$. *This corresponds to* $\alpha_\ell(3,6) = 1.0705, 1.01968, 0.986506, 0.972334, 0.96299, 0.95742, 0.953647$. *We conclude that* $\alpha_\infty(3,6) \approx 0.95$.

In general, scaling laws are only guaranteed to give accurate predictions "close" to the threshold and we have no a priori bound on the absolute size of the error that one gets for a fixed size. Figure 2.2 compares the average block error probability $\mathbb{E}_{\mathtt{G}}[P_B(\mathtt{G}, \epsilon)]$ determined via simulations with the performance predicted via scaling for the $(3,4)$ as well as the $(3,6)$-regular ensemble. As we see, in both cases the predictions are very good indicators of the performance of the code. As one would expect, as the block length increases, the predictions become more and more accurate.

2.5 Flipping Probabilities

Since we are working at the threshold, x_ℓ and $y_{\ell+1}$ are constant (by definition of the threshold) for all $\ell \geq 1$. So, in order to simplify the notation, we drop the iteration indices and use x and y in the following.

As a first ingredient we need to know the joint conditional probability that a variable-to-check message is in error/correct at time ℓ and in error/correct at time $\ell - j$ given that its received value is in error/correct. Note that these quantities are more involved than the equivalent quantities for the BEC: for the BEC messages are monotone – either they will become known at some point in time and stay so, or they are permanently the erasure message. But there are no equivalent monotonicity property for the present case. The fact that even on an infinite trees, above the threshold, messages in general do not converge but continue to "flip" is a main new ingredient when analyzing general message passing algorithms (compared to the BEC case). It makes the analysis considerably more difficult.

Lemma 2.2. *Consider a variable node* vn *and its outgoing message* v. *The event* $v = E$ *means that message* v *is in error and the event* $v \neq E$ *means that*

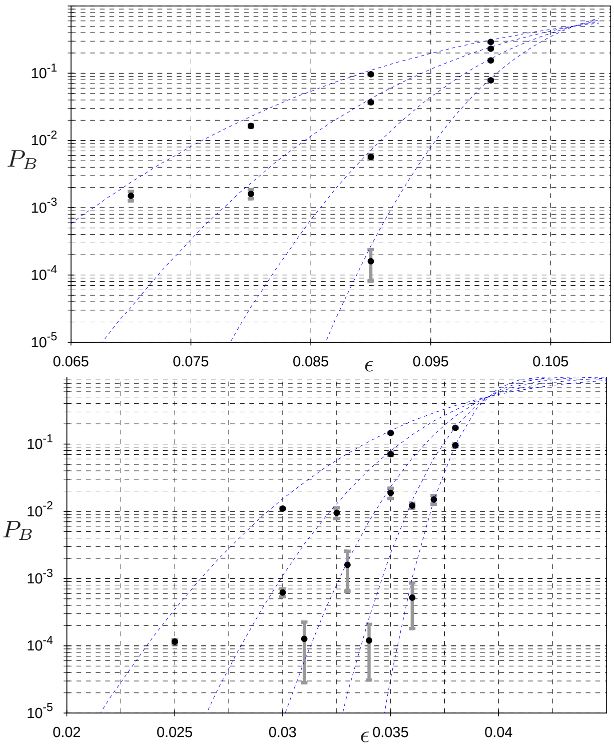

Figure 2.2: Top: Comparison of the block error probability $\mathbb{E}_{\mathtt{G}\in\text{LDPC}(nx^3,\frac{3}{4}nx^4)}[P_B(\mathtt{G},\epsilon)]$ (dots with 95% confidence intervals) determined via simulations with the prediction given by the scaling law. The lengths are $n = 512$, 1024, 2048, and 4096, respectively. Since the scaling law predicts only large-sized errors, the ensembles were expurgated: only error events of size at least 70, 100, 200, respectively, were counted. Bottom: Comparison of the block error probability $\mathbb{E}_{\mathtt{G}\in\text{LDPC}(nx^3,\frac{1}{2}nx^6)}[P_B(\mathtt{G},\epsilon)]$ (solid line) determined via simulations with the prediction given by the scaling law (dashed line). The lengths are $n = 2048$, 4096, 8192, 16384 and 32768 respectively. Since the scaling law predicts only large-sized errors, the ensembles were expurgated: only error events of size at least 50, 70, 100, 200 and 200, respectively, were counted.

message v is correct. Let us define

$$\begin{aligned}
p_{c,c|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell \neq E, v_{\ell-j} \neq E|\mathrm{vn} \neq E)\\
p_{c,e|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell \neq E, v_{\ell-j} = E|\mathrm{vn} \neq E)\\
p_{e,c|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell = E, v_{\ell-j} \neq E|\mathrm{vn} \neq E)\\
p_{e,e|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell = E, v_{\ell-j} = E|\mathrm{vn} \neq E)\\
p_{c,c|e}(\ell,\ell-j) &:= \mathbb{P}(v_\ell \neq E, v_{\ell-j} \neq E|\mathrm{vn} = E)\\
p_{c,e|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell \neq E, v_{\ell-j} = E|\mathrm{vn} = E)\\
p_{e,c|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell = E, v_{\ell-j} \neq E|\mathrm{vn} = E)\\
p_{e,e|c}(\ell,\ell-j) &:= \mathbb{P}(v_\ell = E, v_{\ell-j} = E|\mathrm{vn} = E)
\end{aligned}$$

Then

$$\begin{aligned}
p_{c,c|c}(\ell,\ell) &= 1 - y^{\mathtt{l}-1} & p_{c,e|c}(\ell,\ell) &= 0\\
p_{e,c|c}(\ell,\ell) &= 0 & p_{e,e|c}(\ell,\ell) &= y^{\mathtt{l}-1}\\
p_{c,c|e}(\ell,\ell) &= \bar{y}^{\mathtt{l}-1} & p_{c,e|e}(\ell,\ell) &= 0\\
p_{e,c|e}(\ell,\ell) &= 0 & p_{e,e|e}(\ell,\ell) &= 1 - \bar{y}^{\mathtt{l}-1}\\
p_{c,c|c}(\ell,0) &= 1 - y^{\mathtt{l}-1} & p_{c,e|c}(\ell,0) &= 0\\
p_{e,c|c}(\ell,0) &= y^{\mathtt{l}-1} & p_{e,e|c}(\ell,0) &= 0,\\
p_{c,c|e}(\ell,0) &= 0 & p_{c,e|e}(\ell,0) &= \bar{y}^{\mathtt{l}-1}\\
p_{e,c|e}(\ell,0) &= 0 & p_{e,e|e}(\ell,0) &= 1 - \bar{y}^{\mathtt{l}-1}.
\end{aligned}$$

$$\begin{aligned}
p_{c,c|c}(\ell,\ell-j) &= 1 - 2y^{\mathtt{l}-1} + q_{e,e}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{c,e|c}(\ell,\ell-j) &= y^{\mathtt{l}-1} - q_{e,e}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{e,c|c}(\ell,\ell-j) &= y^{\mathtt{l}-1} - q_{e,e}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{e,e|c}(\ell,\ell-j) &= q_{e,e}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{c,c|e}(\ell,\ell-j) &= q_{c,c}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{c,e|c}(\ell,\ell-j) &= \bar{y}^{\mathtt{l}-1} - q_{c,c}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{e,c|c}(\ell,\ell-j) &= \bar{y}^{\mathtt{l}-1} - q_{c,c}(\ell,\ell-j)^{\mathtt{l}-1}\\
p_{e,e|c}(\ell,\ell-j) &= 1 - 2\bar{y}^{\mathtt{l}-1} + q_{c,c}(\ell,\ell-j)^{\mathtt{l}-1}
\end{aligned}$$

where

$$
\begin{aligned}
&q_{a,b}(\ell, \ell - j) \\
&= \frac{1}{4}\left((p_{c,c}(\ell, \ell - j) + p_{c,e}(\ell, \ell - j) + \right. \\
&\qquad p_{e,c}(\ell, \ell - j) + p_{e,e}(\ell, \ell - j))^{\mathtt{r}-1} \\
&\quad (-1)^{1_{\{a=e\}}} (p_{c,c}(\ell, \ell - j) - p_{c,e}(\ell, \ell - j) \\
&\qquad + p_{e,c}(\ell, \ell - j) - p_{e,e}(\ell, \ell - j))^{\mathtt{r}-1} \\
&\quad (-1)^{1_{\{b=e\}}} + (p_{c,c}(\ell, \ell - j) + p_{c,e}(\ell, \ell - j) \\
&\qquad - p_{e,c}(\ell, \ell - j) - p_{e,e}(\ell, \ell - j))^{\mathtt{r}-1} \\
&\quad (-1)^{(1_{\{a=e\}} + 1_{\{b=e\}})} (p_{c,c}(\ell, \ell - j) - p_{c,e}(\ell, \ell - j) \\
&\qquad \left. - p_{e,c}(\ell, \ell - j) + p_{e,e}(\ell, \ell - j))^{\mathtt{r}-1}\right)
\end{aligned}
$$

with

$$
\begin{aligned}
p_{c,c}(\ell, \ell - j) &= \bar{\epsilon} p_{c,c|c}(\ell, \ell - j) + \epsilon p_{c,c|e}(\ell, \ell - j) \\
p_{c,e}(\ell, \ell - j) &= \bar{\epsilon} p_{c,e|c}(\ell, \ell - j) + \epsilon p_{c,e|e}(\ell, \ell - j) \\
p_{e,c}(\ell, \ell - j) &= \bar{\epsilon} p_{e,c|c}(\ell, \ell - j) + \epsilon p_{e,c|e}(\ell, \ell - j) \\
p_{e,e}(\ell, \ell - j) &= \bar{\epsilon} p_{e,e|c}(\ell, \ell - j) + \epsilon p_{e,e|e}(\ell, \ell - j)
\end{aligned}
$$

2.6 Variance Computation

Considering an edge connected to a variable node vn with its variable-to-check message v. We associate to this edge a new value

$$
F_\ell = \begin{cases} 0_\ell \text{ if } \mathrm{vn} \neq E, v_\ell = E, \\ 1_\ell \text{ if } \mathrm{vn} = E, v_\ell \neq E. \end{cases}
$$

Let us label all edges in the graph from 1 to $n\mathtt{l}$. Consider an edge i and let us define $\chi_{i,\ell}$

$$
\chi_{i,\ell} = \begin{cases} 1 \text{ if } F_{i,\ell} = 0_{i,\ell}, \\ -1 \text{ if } F_{i,\ell} = 1_{i,\ell}. \end{cases}
$$

By noting that $X_\ell - X_0 = \sum_i \chi_{i,\ell}$, we can compute the variance at the threshold for $\ell \geq 1$:

$$
\begin{aligned}
\sigma_\ell^2(\epsilon^{\mathrm{BP}}) &= \lim_{n\to\infty} \mathbb{E}[(Z_\ell(\epsilon^{\mathrm{BP}}))^2] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\left(\frac{\sum_i \chi_{i,\ell}}{\sqrt{n\mathtt{l}}}\right)^2\right] \\
&= \lim_{n\to\infty} \frac{\sum_{i,j} \mathbb{E}\left[\chi_{i,\ell}\cdot\chi_{j,\ell}\right]}{n\mathtt{l}} \\
&= \lim_{n\to\infty} \frac{n\mathtt{l}\sum_i \mathbb{E}\left[\chi_{1,\ell}\cdot\chi_{i,\ell}\right]}{n\mathtt{l}} \\
&= \lim_{n\to\infty} \sum_i \mathbb{E}\left[\chi_{1,\ell}\cdot\chi_{i,\ell}\right] \\
&= \bar{\epsilon}y^{\mathtt{l}-1} + \epsilon\bar{y}^{\mathtt{l}-1} \\
&\quad +(\mathtt{l}-1)(\bar{\epsilon}y^{\mathtt{l}} + \epsilon\bar{y}^{\mathtt{l}}) \\
&\quad +\sum_{j=1}^{\ell}\gamma^j\left(p_{\mathtt{T}_1}(0_{0,\ell},0_{j,\ell}) - p_{\mathtt{T}_1}(0_{0,\ell},1_{j,\ell})\right. \\
&\qquad\qquad \left.-p_{\mathtt{T}_1}(1_{0,\ell},0_{j,\ell}) + p_{\mathtt{T}_1}(1_{0,\ell},1_{j,\ell})\right) \\
&\quad +\sum_{j=1}^{\ell}\gamma^j\left(p_{\mathtt{T}_3}(0_{0,\ell},0_{j,\ell}) - p_{\mathtt{T}_3}(0_{0,\ell},1_{j,\ell})\right. \\
&\qquad\qquad \left.-p_{\mathtt{T}_3}(1_{0,\ell},0_{j,\ell}) + p_{\mathtt{T}_3}(1_{0,\ell},1_{j,\ell})\right) \\
&\quad +\sum_{j=1}^{2\ell}(\mathtt{l}-1)\gamma^j\left(p_{\mathtt{T}_4}(0_{0,\ell},0_{j,\ell}) - p_{\mathtt{T}_4}(0_{0,\ell},1_{j,\ell})\right. \\
&\qquad\qquad \left.-p_{\mathtt{T}_4}(1_{0,\ell},0_{j,\ell}) + p_{\mathtt{T}_4}(1_{0,\ell},1_{j,\ell})\right)
\end{aligned}
\tag{2.1}
$$

where

$$\gamma = (\mathtt{l}-1)(\mathtt{r}-1)$$

$$
\begin{aligned}
&p_{\mathtt{T}_1}(0_{0,\ell},0_{j,\ell}) \\
&= (1,0,0,0)\mathrm{M}^j(1,0,0,0)^T\cdot\bar{\epsilon}p_{e,e|c}(\ell,\ell-j) \\
&\quad +(1,0,0,0)\mathrm{M}^j(0,1,0,0)^T\cdot\bar{\epsilon}p_{e,c|c}(\ell,\ell-j)
\end{aligned}
$$

$$
\begin{aligned}
&p_{\mathtt{T}_1}(0_{0,\ell},1_{j,\ell}) \\
&= (1,0,0,0)\mathrm{M}^j(0,0,1,0)^T\cdot\epsilon p_{c,c|e}(\ell,\ell-j) \\
&\quad +(1,0,0,0)\mathrm{M}^j(0,0,0,1)^T\cdot\epsilon p_{c,e|e}(\ell,\ell-j)
\end{aligned}
$$

$$
\begin{aligned}
&p_{\mathtt{T}_1}(1_{0,\ell},0_{j,\ell}) \\
&= (0,0,1,0)\mathrm{M}^j(1,0,0,0)^T\cdot\bar{\epsilon}p_{e,e|c}(\ell,\ell-j) \\
&\quad +(0,0,1,0)\mathrm{M}^j(0,1,0,0)^T\cdot\bar{\epsilon}p_{e,c|c}(\ell,\ell-j)
\end{aligned}
$$

$$\begin{aligned}
&p_{\mathtt{T}_1}(1_{0,\ell}, 1_{j,\ell}) \\
&= (0,0,1,0)\mathrm{M}^j(0,0,1,0)^T \cdot \epsilon p_{c,c|e}(\ell, \ell-j) \\
&\quad + (0,0,1,0)\mathrm{M}^j(0,0,0,1)^T \cdot \epsilon p_{c,e|e}(\ell, \ell-j)
\end{aligned}$$

$$\begin{aligned}
p_{\mathtt{T}_3}(0_{0,\ell}, 0_{j,\ell}) &= p_{\mathtt{T}_1}(0_{0,\ell}, 0_{j,\ell}) \\
p_{\mathtt{T}_3}(0_{0,\ell}, 1_{j,\ell}) &= p_{\mathtt{T}_1}(0_{0,\ell}, 1_{j,\ell}) \\
p_{\mathtt{T}_3}(1_{0,\ell}, 0_{j,\ell}) &= p_{\mathtt{T}_1}(1_{0,\ell}, 0_{j,\ell}) \\
p_{\mathtt{T}_3}(1_{0,\ell}, 1_{j,\ell}) &= p_{\mathtt{T}_1}(1_{0,\ell}, 1_{j,\ell})
\end{aligned}$$

If $j < \ell$:

$$\begin{aligned}
&p_{\mathtt{T}_4}(0_{0,\ell}, 0_{j,\ell}) \\
&= (1,0,0,0)\mathrm{A}_0\mathrm{B}_j(0) \\
&\quad \cdot \prod_{k=1}^{j-1} \mathrm{A}_j(k)\mathrm{B}_j(k) \begin{pmatrix} \bar{\epsilon} y f_{e,e}(\ell, \ell-j) \\ \bar{\epsilon}(y^{1-2}\bar{y} + f_{e,\bar{e}}(\ell, \ell-j)y) \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(0_{0,\ell}, 1_{j,\ell}) \\
&= (0,0,1,0)\mathrm{A}_0\mathrm{B}_j(0) \\
&\quad \cdot \prod_{k=1}^{j-1} \mathrm{A}_j(k)\mathrm{B}_j(k) \begin{pmatrix} \bar{\epsilon} y f_{e,e}(\ell, \ell-j) \\ \bar{\epsilon}(y^{1-2}\bar{y} + f_{e,\bar{e}}(\ell, \ell-j)y) \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(1_{0,\ell}, 0_{j,\ell}) \\
&= (1,0,0,0)\mathrm{A}_0\mathrm{B}_j(0) \\
&\quad \cdot \prod_{k=1}^{j-1} \mathrm{A}_j(k)\mathrm{B}_j(k) \begin{pmatrix} 0 \\ 0 \\ \epsilon(\bar{y}^{1-2}y + f_{c,\bar{c}}(\ell, \ell-j)\bar{y}) \\ \epsilon \bar{y} f_{c,c}(\ell, \ell-j) \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(1_{0,\ell}, 1_{j,\ell}) \\
&= (0,0,1,0)\mathrm{A}_0\mathrm{B}_j(0) \\
&\quad \cdot \prod_{k=1}^{j-1} \mathrm{A}_j(k)\mathrm{B}_j(k) \begin{pmatrix} 0 \\ 0 \\ \epsilon(\bar{y}^{1-2}y + f_{c,\bar{c}}(\ell, \ell-j)\bar{y}) \\ \epsilon \bar{y} f_{c,c}(\ell, \ell-j) \end{pmatrix}
\end{aligned}$$

If $\ell \le j < 2\ell$:

$$\begin{aligned}
&p_{\mathtt{T}_4}(0_{0,\ell}, 0_{j,\ell}) \\
&= (1,0,0,0)\mathrm{M}^{j-\ell}\tilde{\mathrm{A}}_0\mathrm{B}_j(j-\ell) \\
&\quad \cdot \prod_{k=j-\ell+1}^{\ell-1} \mathrm{A}_j(k)\mathrm{B}_j(k)\tilde{\mathrm{A}}_1 \begin{pmatrix} (1,0,0,0)\mathrm{M}^{j-\ell}(1,0,0,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,0,1,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(0_{0,\ell}, 1_{j,\ell}) \\
&= (0,0,1,0)\mathrm{M}^{j-\ell}\tilde{\mathrm{A}}_0\mathrm{B}_j(j-\ell) \\
&\quad \cdot \prod_{k=j-\ell+1}^{\ell-1} \mathrm{A}_j(k)\mathrm{B}_j(k)\tilde{\mathrm{A}}_1 \begin{pmatrix} (1,0,0,0)\mathrm{M}^{j-\ell}(1,0,0,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,0,1,0)^T \\ (1,0,0,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(1_{0,\ell}, 0_{j,\ell}) \\
&= (1,0,0,0)\mathrm{M}^{j-\ell}\tilde{\mathrm{A}}_0\mathrm{B}_j(j-\ell) \\
&\quad \cdot \prod_{k=j-\ell+1}^{\ell-1} \mathrm{A}_j(k)\mathrm{B}_j(k)\tilde{\mathrm{A}}_1 \begin{pmatrix} (0,0,1,0)\mathrm{M}^{j-\ell}(1,0,0,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,0,1,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&p_{\mathtt{T}_4}(1_{0,\ell}, 1_{j,\ell}) \\
&= (0,0,1,0)\mathrm{M}^{j-\ell}\tilde{\mathrm{A}}_0\mathrm{B}_j(j-\ell) \\
&\quad \cdot \prod_{k=j-\ell+1}^{\ell-1} \mathrm{A}_j(k)\mathrm{B}_j(k)\tilde{\mathrm{A}}_1 \begin{pmatrix} (0,0,1,0)\mathrm{M}^{j-\ell}(1,0,0,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,0,1,0)^T \\ (0,0,1,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix}
\end{aligned}$$

If $j = 2\ell$:

$$\begin{aligned}
p_{\mathtt{T}_4}(0_{0,\ell}, 0_{j,\ell}) &= (1,0,0,0)\mathrm{M}^{j-\ell} \begin{pmatrix} 0 \\ \bar{\epsilon}(1,0,0,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ 0 \\ \epsilon(1,0,0,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix} \\
p_{\mathtt{T}_4}(0_{0,\ell}, 1_{j,\ell}) &= (0,0,1,0)\mathrm{M}^{j-\ell} \begin{pmatrix} 0 \\ \bar{\epsilon}(1,0,0,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ 0 \\ \epsilon(1,0,0,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix} \\
p_{\mathtt{T}_4}(1_{0,\ell}, 0_{j,\ell}) &= (1,0,0,0)\mathrm{M}^{j-\ell} \begin{pmatrix} 0 \\ \bar{\epsilon}(0,0,1,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ 0 \\ \epsilon(0,0,1,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix} \\
p_{\mathtt{T}_4}(1_{0,\ell}, 1_{j,\ell}) &= (0,0,1,0)\mathrm{M}^{j-\ell} \begin{pmatrix} 0 \\ \bar{\epsilon}(0,0,1,0)\mathrm{M}^{j-\ell}(0,1,0,0)^T \\ 0 \\ \epsilon(0,0,1,0)\mathrm{M}^{j-\ell}(0,0,0,1)^T \end{pmatrix}
\end{aligned}$$

and

$$\mathrm{M} = \begin{pmatrix} \bar{\epsilon}y^{1-2}g_e & \bar{\epsilon}y^{1-2}g_o & \bar{\epsilon}y^{1-2}g_o & \bar{\epsilon}y^{1-2}g_e \\ \bar{\epsilon}(1-y^{1-2}g_e) & \bar{\epsilon}(1-y^{1-2}g_o) & \bar{\epsilon}(1-y^{1-2}g_o) & \bar{\epsilon}(1-y^{1-2}g_e) \\ \epsilon\bar{y}^{1-2}g_o & \epsilon\bar{y}^{1-2}g_e & \epsilon\bar{y}^{1-2}g_e & \epsilon\bar{y}^{1-2}g_o \\ \epsilon(1-\bar{y}^{1-2}g_o) & \epsilon(1-\bar{y}^{1-2}g_e) & \epsilon(1-\bar{y}^{1-2}g_e) & \epsilon(1-\bar{y}^{1-2}g_o) \end{pmatrix}$$

$$\mathrm{A}_j(k) = \left(\begin{matrix} \epsilon + \bar{\epsilon}f_{e,e}(\ell-k, \ell-j+k) & \epsilon(1-\bar{y}^{1-2}) \\ \bar{\epsilon}f_{\bar{e},e}(\ell-k, \ell-j+k) & \epsilon\bar{y}^{1-2} + \bar{\epsilon}y^{1-2} \\ \epsilon(1-\bar{y}^{1-2}) & \epsilon f_{\bar{c},\bar{c}}(\ell-k, \ell-j+k) \\ 0 & \epsilon f_{c,\bar{c}}(\ell-k, \ell-j+k) \end{matrix} \right.$$

$$\left. \begin{matrix} \bar{\epsilon}f_{e,\bar{e}}(\ell-k, \ell-j+k) & 0 \\ \bar{\epsilon}f_{\bar{e},\bar{e}}(\ell-k, \ell-j+k) & \bar{\epsilon}(1-y^{1-2}) \\ \epsilon\bar{y}^{1-2} + \bar{\epsilon}y^{1-2} & \epsilon f_{\bar{c},c}(\ell-k, \ell-j+k) \\ \bar{\epsilon}(1-y^{1-2}) & \epsilon f_{c,c}(\ell-k, \ell-j+k) + \bar{\epsilon} \end{matrix} \right)$$

$$\mathrm{B}_j(k) = \left(\begin{array}{cc} g_{e,e}(\ell-k-1,\ell-j+k) & g_{o,e}(\ell-k-1,\ell-j+k) \\ g_{o,e}(\ell-k-1,\ell-j+k) & g_{e,e}(\ell-k-1,\ell-j+k) \\ g_{e,o}(\ell-k-1,\ell-j+k) & g_{o,o}(\ell-k-1,\ell-j+k) \\ g_{o,o}(\ell-k-1,\ell-j+k) & g_{e,o}(\ell-k-1,\ell-j+k) \end{array} \right.$$

$$\left. \begin{array}{cc} g_{e,o}(\ell-k-1,\ell-j+k) & g_{o,o}(\ell-k-1,\ell-j+k) \\ g_{o,o}(\ell-k-1,\ell-j+k) & g_{e,o}(\ell-k-1,\ell-j+k) \\ g_{e,e}(\ell-k-1,\ell-j+k) & g_{o,e}(\ell-k-1,\ell-j+k) \\ g_{o,e}(\ell-k-1,\ell-j+k) & g_{e,e}(\ell-k-1,\ell-j+k) \end{array} \right)$$

$$\mathrm{A}_0 = \left(\begin{array}{cc} y\bar{\epsilon}f_{e,e}(\ell,\ell-j) & 0 \\ y\bar{\epsilon}f_{\bar{e},e}(\ell,\ell-j) & y\bar{\epsilon}y^{\mathtt{l}-2} \\ 0 & y\epsilon\bar{y}^{\mathtt{l}-2}+\bar{y}\epsilon f_{c,\bar{c}}(\ell,\ell-j) \\ y\epsilon+\bar{y}\epsilon(1-\bar{y}^{\mathtt{l}-2}) & y\epsilon(1-\bar{y}^{\mathtt{l}-2})+\bar{y}\epsilon f_{\bar{c},\bar{c}}(\ell,\ell-j) \end{array} \right.$$

$$\left. \begin{array}{cc} y\bar{\epsilon}f_{e,\bar{e}}(\ell,\ell-j)+\bar{y}\bar{\epsilon}y^{\mathtt{l}-2} & 0 \\ y\bar{\epsilon}f_{\bar{e},\bar{e}}(\ell,\ell-j)+\bar{y}\bar{\epsilon}(1-y^{\mathtt{l}-2}) & y\bar{\epsilon}(1-y^{\mathtt{l}-2})+\bar{y}\bar{\epsilon} \\ 0 & \bar{y}\epsilon f_{c,c}(\ell,\ell-j) \\ \bar{y}\epsilon\bar{y}^{\mathtt{l}-2} & \bar{y}\epsilon f_{\bar{c},c}(\ell,\ell-j) \end{array} \right)$$

$$\tilde{\mathrm{A}}_0 = \begin{pmatrix} 0 & 0 & \bar{\epsilon}y^{\mathtt{l}-2} & 0 \\ 0 & 0 & \bar{\epsilon}(1-y^{\mathtt{l}-2}) & \bar{\epsilon} \\ 0 & \epsilon\bar{y}^{\mathtt{l}-2} & 0 & 0 \\ \epsilon & \epsilon(1-\bar{y}^{\mathtt{l}-2}) & 0 & 0 \end{pmatrix}$$

$$\tilde{\mathrm{A}}_1 = \begin{pmatrix} 0 & 0 & 0 & \epsilon \\ \bar{\epsilon}y^{\mathtt{l}-2} & \bar{\epsilon}(1-y^{\mathtt{l}-2}) & 0 & 0 \\ 0 & 0 & \epsilon\bar{y}^{\mathtt{l}-2} & \epsilon(1-\bar{y}^{\mathtt{l}-2}) \\ 0 & \bar{\epsilon} & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} f_{e,e}(\ell,\ell-j) &= q_{e,e}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{e,\bar{e}}(\ell,\ell-j) &= y^{\mathtt{l}-2}-q_{e,e}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{\bar{e},e}(\ell,\ell-j) &= y^{\mathtt{l}-2}-q_{e,e}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{\bar{e},\bar{e}}(\ell,\ell-j) &= 1-2y^{\mathtt{l}-2}+q_{e,e}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{c,c}(\ell,\ell-j) &= q_{c,c}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{c,\bar{c}}(\ell,\ell-j) &= \bar{y}^{\mathtt{l}-2}-q_{c,c}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{\bar{c},c}(\ell,\ell-j) &= \bar{y}^{\mathtt{l}-2}-q_{c,c}(\ell,\ell-j)^{\mathtt{l}-2} \\ f_{\bar{c},\bar{c}}(\ell,\ell-j) &= 1-2\bar{y}^{\mathtt{l}-2}+q_{c,c}(\ell,\ell-j)^{\mathtt{l}-2} \end{aligned}$$

$$g_e = \frac{1+(1-2x)^{\mathtt{r}-2}}{2}, \quad g_o = \frac{1-(1-2x)^{\mathtt{r}-2}}{2}$$

$$\begin{aligned} & g_{e,e}(\ell,\ell-j) \\ &= \frac{1}{4}\Big((p_{c,c}(\ell,\ell-j)+p_{c,e}(\ell,\ell-j) \\ &\qquad +p_{e,c}(\ell,\ell-j)+p_{c,c}(\ell,\ell-j))^{\mathtt{r}-2} \\ &+(p_{c,c}(\ell,\ell-j)-p_{c,e}(\ell,\ell-j) \\ &\qquad +p_{e,c}(\ell,\ell-j)-p_{e,e}(\ell,\ell-j))^{\mathtt{r}-2} \\ &+(p_{c,c}(\ell,\ell-j)+p_{c,e}(\ell,\ell-j) \\ &\qquad -p_{e,c}(\ell,\ell-j)-p_{e,e}(\ell,\ell-j))^{\mathtt{r}-2} \\ &+(p_{c,c}(\ell,\ell-j)-p_{c,e}(\ell,\ell-j) \\ &\qquad -p_{e,c}(\ell,\ell-j)+p_{e,e}(\ell,\ell-j))^{\mathtt{r}-2}\Big) \end{aligned}$$

$$\begin{aligned}
g_{e,o}(\ell,\ell-j) &= g_e - g_{e,e}(\ell,\ell-j) \\
g_{o,e}(\ell,\ell-j) &= g_e - g_{e,e}(\ell,\ell-j) \\
g_{o,o}(\ell,\ell-j) &= 1 - 2g_e + g_{e,e}(\ell,\ell-j)
\end{aligned}$$

2.7 Summary and Open Problems

We investigated the scaling law for the Gallager algorithm A. More precisely, we considered ensembles whose threshold is determined by a fixed point right at the beginning of the decoding process. We saw how to derive the scaling law and how to compute the scaling parameters. By comparing the predictions to simulation results we confirmed that these predictions are very accurate already for moderate lengths. Two major new ingredients entered in the analysis compared to the analysis for the BEC. First, for the present the fixed point was located at the beginning of the decoding process. This is somewhat specific to the Gallager A algorithm and does not seem to be the typical case for message passing algorithms with higher quantization. The second ingredient though seems typical for the general case. When we consider density evolution above the threshold on an infinite tree then the *desities* of the messages converge but the actual messages do not. This "flipping" of the messages introduces a significant new component which makes the analysis both, more interesting but also more challenging.

Many interesting questions remain open. One could investigate the scaling of ensembles whose threshold is given by a fixed point in the middle of the decoding process or the scaling of ensembles whose threshold is determined by the stability condition. The latter one is likely to be of a different nature than the preceding two. In order to use these scaling laws in a context of optimization it is also necessary to accomplish the variance computations for the general irregular case and to determine the error floor under Gallager A. What are the *pseudocodewords* of this decoder?

3 General Scaling Law

3.1 EXIT-Like Curves

Consider transmission over a BMS channel decoded with a generic quantized message-passing (QMP) decoder which satisfies the standard message-passing symmetry conditions [12]. To be concrete, assume that the messages sent in the decoder are quantized and that they take their values in a real-valued finite alphabet $\{-w_m, \ldots, w_m\}$, $m \in \mathbb{N}$, where $0 = w_0 < w_1 < \cdots < w_m$. We do allow the MP decoding rules to be time dependent for a finite number of steps before they settle on a fixed rule. Given the symmetry of the channel and the decoder, we can assume that the all-zero codeword was transmitted.

In order to derive our scaling law we need several ingredients. First, we need to define a suitable "EXIT-like" curve. By this we mean a scalar quantity which characterizes the behavior of the decoder as a function of the channel parameter. Many choices are possible. Let the channel be parameterized by $\mathtt{h}$, the entropy of the channel. For every $\mathtt{h} \in [0, 1]$, initialize the density evolution process with (a suitably quantized version of) the density $\mathsf{a}_{\mathsf{BMS}}(\mathtt{h})$. Let a denote the resulting fixed point of density evolution. (This density is a function of the channel parameter, but to simplify notation we omit this dependence.) More precisely, this is the variable to check node message density. Pick a function $\mathfrak{E}$, $\mathfrak{E} : \mathbb{R}^{2m+1} \mapsto \mathbb{R}$. For the sake of definiteness, let us assume that for a density $\mathsf{v} \in \mathbb{R}^{2m+1}$, $\mathfrak{E}(\mathbf{v}) = 1 - v_m$, where v_m is the value of the maximum component of $\mathbf{v}$. Our language below will reflect this particular choice, but mathematically many other choices are possible and might sometimes be more convenient. E.g., in Section 3.4 it is notationally more convenient to pick $\mathfrak{E}(\mathbf{v}) = v_m$. A futher natural choice would be to pick the number of negative messages (which is equal to the error probability). Generically, all these choices give identical result.

Define $x = \mathfrak{E}(\mathsf{a})$. Note that x "measures" the performance of the decoder.

The special case x^{MP} is the value that x takes on when $\mathtt{h} = \mathtt{h}^{\mathrm{MP}}$, the threshold value. To derive from the family of fixed point an EXIT-like curve we can e.g. plot $(\mathtt{h}, x)$. Our choice of $\mathfrak{C}$ implies that if we are below the threshold, we expect x to be equal to 0 (at the end of the decoding process all messages are positive and take on the highest reliability), whereas above the threshold it will have a non-zero value. This will give us a curve which is somewhat reminiscent of a standard EXIT curve (of the overall code), explaining our choice of name.

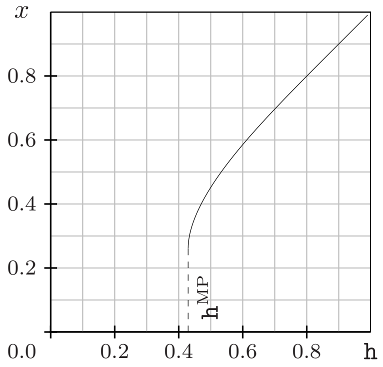

Figure 3.1: EXIT-like curve for the (3,6)-code over the BEC: $\mathtt{h}^{\mathrm{MP}} \approx 0.42944$.

Consider the EXIT-like curve depicted in Figure 3.1. Mathematically it is given by

$$x(\mathtt{h}) = \begin{cases} ,(\mathtt{h}, 0), & \mathtt{h} \in [0, \mathtt{h}^{\mathrm{MP}}) \\ (\mathtt{h}(x), x), & \mathtt{h} \in (\mathtt{h}^{\mathrm{MP}}, 1] \leftrightarrow x \in (x^{\mathrm{MP}}, 1]. \end{cases} \tag{3.1}$$

In order for an EXIT-like curve to be suitable for our derivation, it must have the following property: at the threshold, the curve must have an infinite slope and the second derivative must be strictly non-zero. One can glean from the picture that this is true for the specific curve shown. We will say that an EXIT-like curve is *regular* if it has the above property.

3.2 General Derivation of Scaling Law

We will proceed in the following way. We first rederive the scaling law for the BEC using the EXIT-like curve instead of the density evolution curve associated to the peeling decoder. Although the derivation will depend on the specific channel and decoder which we assume we will see that the final expression is meaningful in a much broader context. This will then be the starting point for the general scaling law.

Therefore, consider transmission over the BEC. Imagine the following experiment. We start with a randomly chosen graph from the ensemble and all bits erased. Choose a random bit and reveal it to the decoder. Next run the iterative decoding process until it is stuck and the decoder hits a fixed point.

This gives rise to one fixed point pair (H, X), where H equals the number of not yet revealed variables and where X is equal to the number of still erased variable-to-check messages. We continue now in this fashion, each time choosing at random a new bit which we then reveal and running the decoder until it is stuck. This gives us a sequence of fixed point pairs $(H^{(t)}, X^{(t)})$. Let us denote $H^{(t)}$ normalized to n, the length of the ensemble, as $\hat{H}^{(t)} = H^{(t)}/n$ and $X^{(t)}$ normalized to $n\Lambda'(1)$, the number of edges, as $\hat{X}^{(t)} = X^{(t)}/n\Lambda'(1)$. If we connect the points $(\hat{H}^{(t)}, \hat{X}^{(t)})$, then we get a "curve".

Such a curve is of course random, but if we increase the blocklength, then we can observe a concentration of the individual instances around an "average curve". The exact description of this average curve is easy to write down in terms of density evolution quantities.

Let $\hat{X}_{\mathtt{h}}$ denote the random variable which we get if we look at $\hat{X}^{(t)}$ at the time when $\hat{H}^{(t)} = \mathtt{h}$. In a similar way, let $\hat{H}_x$ denote the random variable which we get if we look at $\hat{H}^{(t)}$ at the *first* time when $\hat{X}^{(t)} = x$. Imagine that we are sitting just above the threshold at $\mathtt{h} = \mathtt{h}^{\text{MP}} + \Delta\mathtt{h}$. Due to the assumption that the EXIT-like curve is regular, this corresponds to the parameter $x = x^{\text{MP}} + \sqrt{\Delta\mathtt{h}}c$, where c is a constant. Note that $\mathbb{E}[\hat{H}_x] = \mathtt{h}^{\text{MP}} + \Delta\mathtt{h}$. Let $\sigma_{\mathtt{h}}^2 = \mathbb{E}[(\hat{H}_x - \mathbb{E}[\hat{H}_x])^2]$. Note that $\sigma_{\mathtt{h}}$ is a continuous function of the channel parameter so that we can consider $\sigma_{\mathtt{h}^{\text{MP}}} = \lim_{\mathtt{h}\to\mathtt{h}^{\text{MP}}} \sigma_{\mathtt{h}}$. Consider again the decoding process as described above, which gives rise to the sequence of random variables $(H^{(t)}, X^{(t)})$. Assume that we measure the X-component of this process for a particular instance t. We claim that if $\hat{X}^{(t)} < x^{\text{MP}}$ then with high probability the decoder for this instance will finish without revealing any further bits. On the other hand if $\hat{X}^{(t)} > x^{\text{MP}}$ then with high probability further revelations of bits will be necessary. This is due to the shape of the EXIT-like curve: around the point x^{MP} the curve has an infinite slope, indicating that the decoder is at the brink of successful decoding with the given information. Hence, if $\hat{H}_{x^{\text{MP}}} < \mathtt{h}$, then the given instance would not have decoded successfully with high probability if we had transmitted over a channel with parameter $\mathtt{h}$, and conversely it would have been successful with high probability if $\hat{H}_{x^{\text{MP}}} > \mathtt{h}$. Our probability of error estimate is therefore

$$P_B(\mathtt{h}) = \mathbb{P}\{\hat{H}_{x^{\text{MP}}} < \mathtt{h}\} = Q\left(\frac{\Delta\mathtt{h}}{\sigma_{\mathtt{h}^{\text{MP}}}}\right).$$

In the last step we have assumed that the distribution of $\hat{H}_{x^{\text{MP}}}$ is Gaussian with mean $\mathtt{h}^{\text{MP}}$ and standard deviation $\sigma_{\mathtt{h}^{\text{MP}}}$. This was shown to hold in the original derivation in [4].

In general it is not an easy task to compute $\sigma_{\mathtt{h}^{\text{MP}}}$ directly. But we can relate it to the variance of the messages which is amenable to an analytic analysis in the following way. Due to the regularity of the EXIT-like curve, we have for $\Delta\mathtt{h}$ sufficiently small

$$\Delta\mathtt{h} = \mathtt{h}'(x)\Delta x = \mathtt{h}'(x)\frac{X_{\mathtt{h}} - \mathbb{E}[X_{\mathtt{h}}]}{n\Lambda'(1)}.$$

If we take the square and the expectation on both sides we obtain

$$\sigma_{\mathtt{h}}^2 = \frac{(\mathtt{h}'(x))^2\mathcal{V}}{n\Lambda'(1)},$$

where $\mathcal{V} = \lim_{n\to\infty} \frac{\mathbb{E}[(X_{\mathtt{h}}-\mathbb{E}[X_{\mathtt{h}}])^2]}{n\Lambda'(1)}$. If we take the limit $x \downarrow x^{\text{MP}}$, we thus have

$$\begin{aligned}
\sigma_{\mathtt{h}}^2 \mid_{\mathtt{h}=\mathtt{h}^{\text{MP}}} &= \lim_{x\downarrow x^{\text{MP}}} \frac{(\mathtt{h}'(x))^2\mathcal{V}}{n\Lambda'(1)} \\
&= \frac{1}{n\Lambda'(1)}\left(\frac{\partial^2 \mathtt{h}(x)}{\partial x^2}\Big|_{x=x^{\text{MP}}}\right)^2 \lim_{x\downarrow x^{\text{MP}}} (x-x^{\text{MP}})^2\mathcal{V} \\
&= \frac{1}{n\Lambda'(1)}\left(\frac{\partial^2 \mathtt{h}(x)}{\partial x^2}\Big|_{x=x^{\text{MP}}}\right)^2 \lim_{x\downarrow x^{\text{MP}}} \left(\frac{(x-x^{\text{MP}})}{1-\lambda_2(x)}\right)^2 \xi.
\end{aligned}$$

The last step warrants some remarks. When $h \downarrow h^{\text{MP}}$, i.e., when the channel parameter approaches the threshold from above, then the variance $\mathcal{V}$ diverges. Indeed $\mathcal{V}$ has the form

$$\frac{\xi}{(1-\lambda_2(x))^2} + O\left(\frac{1}{1-\lambda_2(x)}\right),$$

where $\lambda_2(x)$ is a function so that $\lim_{x\downarrow x^{\text{MP}}} \lambda_2(x) = 1$. In fact, λ_2 is the second largest eigenvalue of a matrix related to the density evolution process.

As we remarked at the start of this section, although the derivation was based assuming that transmission takes place over the BEC, the final expression we have derived can be meaningfully be interpreted in a general setting. We are therefore now able to conjecture the following scaling law.

Conjecture 3.1. *Consider a BMS(*$\mathtt{h}$*) channel and an EXIT-like curve which is* regular. *Then*

$$P_B = Q\left(\frac{\sqrt{n}(\mathtt{h}-\mathtt{h}^{\text{MP}}-\beta n^{-\frac{2}{3}})}{\alpha}\right) + (1+o(1)), \tag{3.2}$$

where $\alpha = \frac{\partial^2 \mathtt{h}(x)}{\partial x^2}\big|_{x=x^{\text{MP}}} \lim_{x\downarrow x^{\text{MP}}} \frac{(x-x^{\text{MP}})}{1-\lambda_2(x)}\sqrt{\frac{\xi}{\Lambda'(1)}}$ *and* β *is the shift parameter of the threshold that we mentioned earlier*[1]*.*

Therefore, in order to compute α, we need to compute

$$\frac{\partial^2 h(x)}{\partial x^2}\Big|_{x=x^{\text{MP}}}, \lim_{x\downarrow x^{\text{MP}}} \frac{(x-x^{\text{MP}})}{1-\lambda_2(x)},$$

[1] In these notes, we do not give further details about the parameter β. Nevertheless, let us mention that Amraoui et al. conjectured its expression for the BEC case in [8]. As for α, the computation of β in [8] is based on the peeling decoder behavior. A new approach to compute β in the general case has been proposed in [13]. Finally, Dembo and Montanari proved the scaling law in the case of Poisson ensembles in [14].

as well as ξ. The most difficult quantity is the last one. The computation of ξ has been accomplished for *regular* ensembles in [9]. Therefore, for regular ensembles the scaling parameter α can be computed and can be used to predict the performance. In the left-hand graph of Figure 3.2, we compare the prediction given by the scaling law, where $\alpha = 0.93$, for a $(3, 4)$-regular code used over a binary additive white Gaussian noise (BAWGN) channel and decoded with a quantized BP decoder with simulation points. We see that there is a good aggrement between the prediction and the simulation.

The aim of the remainder of this chapter is twofold. First, in order to be of practical use, we want to compute the parameter α in the case of irregular ensemble. In this case, as we will see later, the computation of ξ is more challenging. Second, we want to present a detailed account of all nessary computations, since such an account has never appeard in the literature.

In order to compute ξ for the irregular case, we proceed in steps. Let us define for each fixed channel parameter $\mathtt{h}$, $\mathcal{V}^{(\ell)} = \lim_{n\to\infty} \mathrm{Var}(X_{\mathtt{h}}^{(\ell)}/\sqrt{n\Lambda'(1)})$, where $X_{\mathtt{h}}^{(\ell)}$ is the number of variable-to-check messages not equal to w_m after ℓ iterations, according to our choice of $\mathfrak{C}$. For each channel parameter unequal to the threshold, the limit $\lim_{\ell\to\infty} \mathcal{V}^{(\ell)}$ exists and is finite. But as we already mentioned, when we let the channel parameter approach the threshold from above, the value of this variance, i.e., $\mathcal{V} = \lim_{x\downarrow x^{\mathrm{MP}}} \lim_{\ell\to\infty} \mathcal{V}^{(\ell)}$ diverges. Thus, as a first step in the computation of ξ, it is useful to compute $\mathcal{V}^{(\ell)}$. This computation is performed in Section 3.4. One can then hope to extract ξ by taking the limit $\lim_{\ell\to\infty} \mathcal{V}^{(\ell)}$ when we approach the threshold from above. We start in Section 3.3 by defining the precise model that we consider as well as the notation and the conventions that we will use.

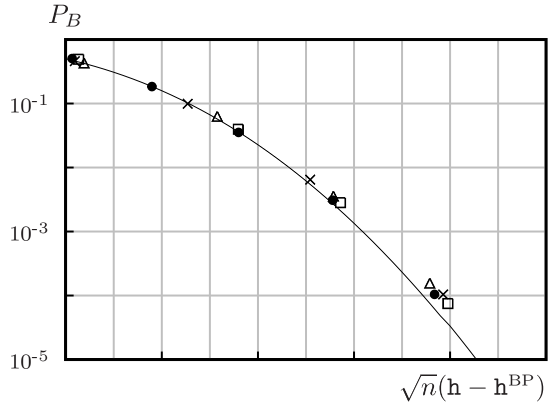

Figure 3.2: Rescaled block error probability with respect to $z = \sqrt{n}(\mathtt{h} - \mathtt{h}^{\mathrm{MP}})$, where $\mathtt{h} = h_2(\sigma)$ and where $h_2(\cdot)$ is the binary entropy function. Note that the simulations points all cluster around the single *mother* curve. This curve is the standard Q-function.

3.3 Notation

Consider transmission over a family of BMS channels characterized by their channel entropy $\mathtt{h}$ and decoded with a QMP decoder that satisfies the standard message-passing symmetry conditions [12]. Further, assume that the messages sent in the decoder are quantized and that they take their values in a real-valued finite alphabet $\mathcal{W} = \{-w_m, \dots, w_m\} = \{-m\Delta, \dots, m\Delta\}$, $m \in \mathbb{N}$, where $\Delta > 0$, denotes the spacing between the members of the alphabet. In the following, in order to simplify the notation, we denote a message value $i\Delta$, by its "position" i. Let us define the operator $\mathfrak{Q} : \mathbb{R} \mapsto \mathcal{W}$, such that $\mathfrak{Q}(s) = \text{sgn}(s)m$, if $|s| \geq m$ and $\lfloor s + 0.5 \rfloor$ otherwise, where the function $\text{sgn}(s)$ corresponds to the sign of s. Let $\Psi : \mathcal{W}^d \mapsto \mathcal{W}$ denote the variable node message map, where d is the node degree. We assume that this map corresponds to the addition of the incoming messages and the quantized channel information followed by a mapping back into the alphabet. More precisely, $\Psi(c, \hat{\nu}_1, \dots, \hat{\nu}_{d-1}) = \mathfrak{Q}(c + \sum_{i=1}^{d-1} \hat{\nu}_i)$, where $\hat{\nu}_i$ are the incoming messages and c is the channel information. Let $\Phi : \mathcal{W}^{d-1} \mapsto \mathcal{W}$ denote the check node message map, where d is the node degree. We assume that the map Φ performes pairwise operations. More precisely $\Phi(\nu_1, \dots, \nu_{d-1}) = \mathfrak{Q}(\dots \mathfrak{Q}(\mathfrak{Q}(\nu_{i_1} \boxplus \nu_{i_2}) \boxplus \nu_{i_3}) \boxplus \cdots \boxplus \nu_{i_{d-1}})$, where ν_i are the incoming messages, $\boxplus$ is the check rule corresponding to the decoder and the indices i_1 to i_{d-1} are assigned randomly from $\{1, \dots, d-1\}$.

Consider a QMP decoder whose processing rules are the ones defined above. In order to define a QMP decoder which closely mimicks a given corresponding non-quantized MP decoder, it is natural implement the processing rules of the non-quantized MP decoder followed by a mapping back into the alphabet. This is what we do at the variable node side, since the addition of two messages is still in a set whose components are equally spaced. On the other hand, the result of $i \boxplus j$, where $i, j \in \mathcal{W}$, does not belong in general to a set whose components are equally spaced. Thus, in order to avoid complexity explosure, it is more convenient to perform the check node rule pairwise.

Since by our assumption both the channel and the decoder are symmetric, the performance of the code is independent of the transmitted codeword. Therefore, let us assume that the all-zero codeword was transmitted. Let $\mathsf{a}^{(\ell)}(\mathtt{h})$ denote the *variable-to-check node* density of density evolution in iteration ℓ assuming that we initialize with the channel density $\mathsf{a}_{\mathsf{BMS}}(\mathtt{h})$. We also define $\mathsf{b}^{(\ell)}(\mathtt{h})$ to be the check-to-variable density of the messages in iteration ℓ. (In the following, in order to simplify notation, we will omit the parameter $\mathtt{h}$ from our notation.) Density evolution is defined as $\mathsf{b}^{(j)} = \sum_d \rho_d (\mathsf{a}^{(j-1)})^{\boxast (d-1)}$, where $\boxast$ denotes the check node operator on a pair of densities followed by the quantization and $\mathsf{a}^{(j)} = \mathsf{a}_{\mathsf{BMS}} \star \sum_d \lambda_d (\mathsf{b}^{(j)})^{\star (d-1)}$, with $\mathsf{a}^{(0)} = \mathsf{a}_{\mathsf{BMS}}$, which corresponds to the average over the ensemble of the convolution of the channel density with $(d-1)$ check-to-variable densities, followed by the quantization.

As explained in the previous section, many choices of the function $\mathfrak{E}$ are possible and give identical result. From now to the end of these notes, we chose $\mathfrak{E}(\mathsf{v}) = v_m$, where $\mathsf{v} \in \mathbb{R}^{2m+1}$. We thus have $x^{(\ell)} = \mathfrak{E}(\mathsf{a}^{(\ell)})$, and the special case x^{MP} is the value of x corresponding to fixed point of the density evolution threshold.

Let us number the edges in the graph from 1 to $n\Lambda'(1)$. To lighten the notation, we define $\mathcal{E} = \{1, \dots n\Lambda'(1)\}$. Given an edge $i \in \mathcal{E}$, we define $\nu_i^{(\ell)}$ to be the variable-to-check message on edge i in iteration ℓ. Similarly, we define $\hat{\nu}_i^{(\ell)}$ to be the check-to-variable message on edge i in iteration ℓ. Finally, let us define $\mu_i^{(\ell)} = \mathbb{1}_{\{\nu_i^{(\ell)}=m\}}$, according to our choice of $\mathfrak{E}$. Note that

$$\mathbb{E}[\mu_i^{(\ell)}] = \mathbb{P}\{\nu_i^{(\ell)} = m\} = \mathfrak{E}(\mathsf{a}^{(\ell)}) = x^{(\ell)}. \tag{3.3}$$

We can then rewrite $\mathcal{V}^{(\ell)}$ as

$$\mathcal{V}^{(\ell)} = \lim_{n\to\infty} \mathrm{Var}\left(\frac{\sum_{i\in\mathcal{E}} \mu_i^{(\ell)}}{\sqrt{n\Lambda'(1)}}\right). \tag{3.4}$$

Example 3.1. [Quantized Belief Propagation] The class of *quantized belief propagation* (QBP) decoders is particularly important. In this case the variable and the check message-passing rules are respectively the maps Ψ and Φ, as defined above. But, in order to have a proper threshold, we need to add the following condition to the check node rule. If we consider two messages ν_1 and ν_2 of highest reliability, the result of $\nu_1 \boxplus \nu_2$ should also be of highest reliability. In other words, $\boxplus$ is the standard BP check node operator to which we add the condition that if $|\nu_1| = |\nu_2| = m$, then $|\nu_1 \boxplus \nu_2| = m$.

3.4 Computation of Variance for Fixed ℓ

Consider $\frac{1}{\sqrt{n\Lambda'(1)}} \sum_{i\in\mathcal{E}} \mu_i^{(\ell)}$. This is a random variable which depends both on the graph and on the noise realization. As stated in (3.4), its variance is equal to $\mathcal{V}^{(\ell)}$. We have

$$\begin{aligned}
\mathcal{V}^{(\ell)} &= \lim_{n\to\infty} \frac{\mathbb{E}\Big[\big(\sum_{i\in\mathcal{E}} \mu_i^{(\ell)}\big)^2\Big] - \mathbb{E}\big[\sum_{i\in\mathcal{E}} \mu_i^{(\ell)}\big]^2}{n\Lambda'(1)} \\
&= \lim_{n\to\infty} \frac{\sum_{j\in\mathcal{E}} \mathbb{E}\big[\sum_{i\in\mathcal{E}} \mu_j^{(\ell)} \mu_i^{(\ell)}\big] - \sum_{j\in\mathcal{E}} \sum_{i\in\mathcal{E}} \mathbb{E}[\mu_j^{(\ell)}]\mathbb{E}[\mu_i^{(\ell)}]}{n\Lambda'(1)} \\
&\overset{\text{(i)}}{=} \lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathcal{E}} \mu_1^{(\ell)} \mu_i^{(\ell)}\Big] - \mathbb{E}[\mu_1^{(\ell)}] \sum_{i\in\mathcal{E}} \mathbb{E}[\mu_i^{(\ell)}] \\
&\overset{(3.3)}{=} \lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathcal{E}} \mu_1^{(\ell)} \mu_i^{(\ell)}\Big] - n\Lambda'(1)(x^{(\ell)})^2,
\end{aligned} \tag{3.5}$$

where, in step (i), we have used the fact that the ensemble is invariant under permutations of its components.

The *computation tree* of $\nu_i^{(\ell)}$, with $i \in \mathcal{E}$, is the tree of height ℓ which has edge i as root and which is formed by all edges whose messages influence the

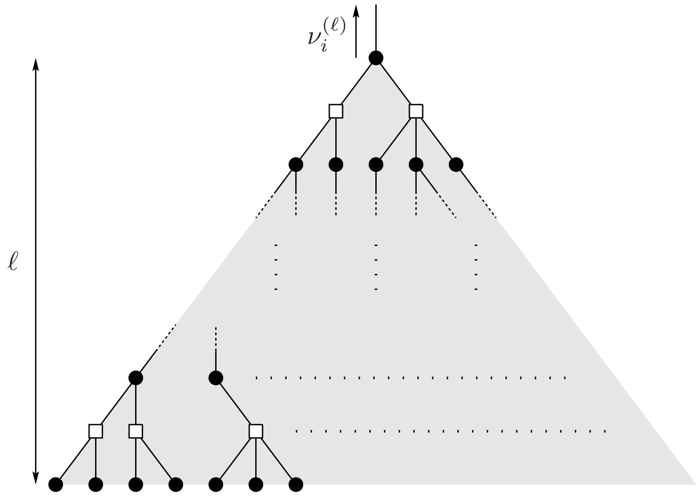

Figure 3.3: Computation tree of $\nu_i^{(\ell)}$.

value of $\nu_i^{(\ell)}$. This is depicted in Figure 3.3. Consider $\nu_1^{(\ell)}$ and let $\mathtt{T}^{(\ell)}$ be the set of indices of all messages whose computation tree intersect the computation tree of $\nu_1^{(\ell)}$, where both trees have depth ℓ. This is shown in Figure 3.4. For convenience, we also add to $\mathtt{T}^{(\ell)}$ indices of edges which are connected to the same variable node as an edge already in $\mathtt{T}^{(\ell)}$. For instance, the edges (b) and (f) in the left-hand side graph of Figure 3.4 are added to $\mathtt{T}^{(\ell)}$, even though their computation trees do not intersect the computation tree of the root edge. Let $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ be the tree formed by all edges which belong to $\mathtt{T}^{(\ell)}$. Thus, $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ is formed by ℓ variable node layers "above" the root variable node and 2ℓ variable nodes layers "below" the root variable, as depicted in the right-hand side of Figure 3.4.

Let us partition $\mathtt{T}^{(\ell)}$ into four sets. Orient all edges in $\mathtt{G}$ from variable node to check node. Let $\mathtt{T}_1^{(\ell)}$ denote the set of edges which are in the "future" as seen by the root edge and which are directed in the same direction as the root edge itself. More precisely, these are the edges within $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ which can be reached by paths starting at the root edge and which point in the same direction as the root edge. Next, let $\mathtt{T}_2^{(\ell)}$ denote the set of edges in the future of the root edge but which point in the opposite direction. Let $\mathtt{T}_3^{(\ell)}$ be the set of edges in the past of the root edge and which point in same direction and, finally, let $\mathtt{T}_4^{(\ell)}$ be the set of edges which are in the past of the root and in which point in the opposite direction. These four types of edges are depicted in the left of Figure

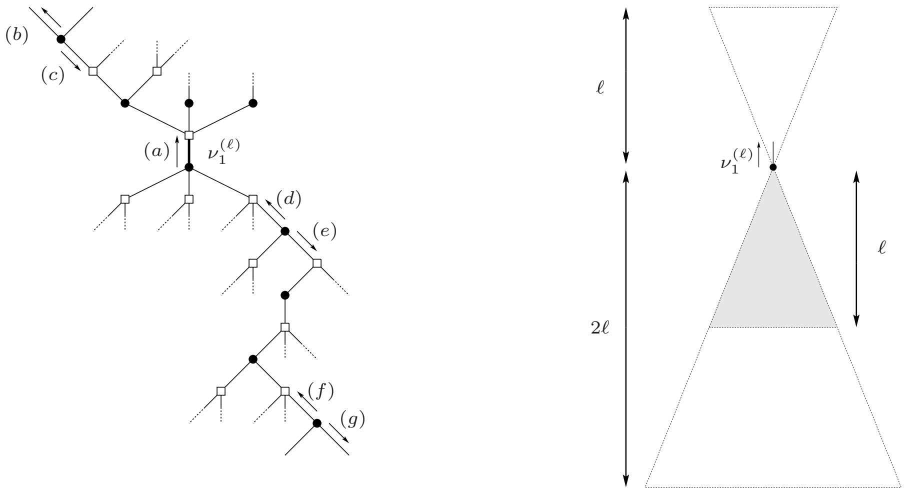

Figure 3.4: *Left:* Graph representing the four types of edges contained in $\mathtt{T}^{(2)}$: (a) root edge, (b) element of $\mathtt{T}_1^{(2)}$, (c) element of $\mathtt{T}_2^{(2)}$, (d) and (f) are elements of $\mathtt{T}_3^{(2)}$ and (e) and (g) are elements of $\mathtt{T}_4^{(2)}$; *Right:* The graph $\mathrm{G}_{\mathtt{T}}^{(\ell)}$. It contains ℓ layers of variable nodes "above" the root node and 2ℓ "below". The gray area represents the computation tree of the root edge.

3.4. Moreover, we define for $i = 1, \dots, 4$, $\mathtt{B}_{i,k}$ to be the set of edges which are in $\mathtt{T}_i^{(\ell)}$ at distance k from the root (which means k levels of variable and check nodes). Finally, let $(\mathtt{T}^{(\ell)})^c$ be the complement of $\mathtt{T}^{(\ell)}$ in $\{1, \dots, n\Lambda'(1)\}$ and let $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ be the graph formed by the edges in $(\mathtt{T}^{(\ell)})^c$. Let us expand (3.5) as

$$\begin{aligned}
\mathcal{V}^{(\ell)} &= \lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}^{(\ell)}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big] + \lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big] - n\Lambda'(1)x_\ell^2 \\
&= \underbrace{\lim_{n\to\infty} \mathbb{E}\big[(\mu_1^{(\ell)})^2\big]}_{x^{(\ell)}} + \underbrace{\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}_1^{(\ell)}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big]}_{\text{edges in } \mathtt{T}_1^{(\ell)}} + \underbrace{\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}_2^{(\ell)}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big]}_{\text{edges in } \mathtt{T}_2^{(\ell)}} \\
&\quad + \underbrace{\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}_3^{(\ell)}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big]}_{\text{edges in } \mathtt{T}_3^{(\ell)}} + \underbrace{\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}_4^{(\ell)}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big]}_{\text{edges in } \mathtt{T}_4^{(\ell)}} \\
&\quad + \underbrace{\lim_{n\to\infty} \Big(\mathbb{E}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big)}_{S^c}. \qquad (3.6)
\end{aligned}$$

The next lemma gives an explicit expression for $\mathcal{V}^{(\ell)}$ in terms of quantities which can be computed via density evolution. Starting from (3.6), we show how we derive this lemma in the Appendices 3.A.1-3.A.5. In Lemma 3.1 we use the following convention for matrix products; we write $\prod_{j=\ell}^{k} \mathrm{M}^{(k)}$ to mean $\mathrm{M}^{(\ell)}\mathrm{M}^{(\ell-1)}\cdots\mathrm{M}^{(k)}$, where $\ell \geq k$ and $\mathrm{M}^{(j)}$ is a matrix depending on j. Note that if $\ell < k$, the product is simply the identity matrix.

Lemma 3.1. [Fixed-ℓ Variance] Consider the irregular LDPC ensemble characterized by its degree distribution $(\lambda(x), \rho(x))$. The variance $\mathcal{V}^{(\ell)}$ is given by

$$
\begin{aligned}
&\mathcal{V}^{(\ell)} = x^{(\ell)} \\
&+ \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbf{e}_m^T \prod_{j=\ell}^{\ell-k+1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{e}_r \mathrm{P}_{m,r}^{(\ell,\ell-k)} \\
&+ \lim_{n\to\infty} (x^{(\ell)})^2 \sum_{k=1}^{\ell} \rho'(1)^k \lambda'(1)^{k-1} \\
&+ \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbf{e}_m^T \prod_{j=\ell}^{\ell-k+1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{e}_r \mathrm{P}_{m,r}^{(\ell,\ell-k)} \\
&+ \lim_{n\to\infty} \sum_{k=\ell+1}^{2\ell} x^{(\ell)} (\rho'(1)\lambda'(1))^{(k-\ell)} \mathbf{e}_m^T \prod_{j=\ell}^{1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)} \mathsf{a}_{\mathsf{BMS}} \\
&+ \lim_{n\to\infty} \sum_{l=0}^{\lceil(\ell-1)/2\rceil} \Big(\prod_{j=0}^{l-1} \hat{\mathrm{B}}^{2l,j}\mathrm{B}^{2l,j+1}\mathbf{e}_s^T \mathsf{b}^{\ell-2l} \mathbb{1}_{\{m=r\}} \Big)^T \\
&\qquad\qquad (\mathrm{F}\mathrm{B}^{2l,0}) \prod_{j=0}^{l-1} \hat{\mathrm{B}}^{2l,j}\mathrm{B}^{2l,j+1}\mathbf{e}_s^T \mathsf{b}^{\ell-2l} \mathbb{1}_{\{m=r\}} \\
&+ \lim_{n\to\infty} \sum_{l=0}^{\lfloor(\ell-1)/2\rfloor} \Big(\prod_{j=0}^{l-1} \hat{\mathrm{B}}^{2l+1,j}\mathrm{B}^{2l+1,j+1}\mathbf{e}_s^T \mathsf{b}^{\ell-2l-1} \mathbb{1}_{\{m=r\}} \Big)^T \\
&\qquad \big(\mathrm{B}^{2l+1,0}\big)^T \mathrm{F}\hat{\mathrm{B}}^{2l+1,-1}\mathrm{B}^{2l+1,0} \prod_{j=0}^{l-1} \hat{\mathrm{B}}^{2l+1,j}\mathrm{B}^{2l+1,j+1}\mathbf{e}_s^T \mathsf{b}^{\ell-2l-1} \mathbb{1}_{\{m=r\}} \\
&+ \lim_{n\to\infty} \sum_{l=\lceil(\ell-1)/2\rceil+1}^{\ell} \Big(\prod_{j=0}^{\ell-l-1} \hat{\mathrm{B}}^{2l,j}\mathrm{B}^{2l,j+1}\mathbf{e}_m^T \prod_{k=\ell}^{2\ell-2l+1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbf{e}_r \mathbb{1}_{\{s=0\}} \Big)^T \\
&\qquad \mathrm{F}\mathrm{B}^{2l,0} \prod_{j=0}^{\ell-l-1} \hat{\mathrm{B}}^{2l,j}\mathrm{B}^{2l,j+1}\mathbf{e}_m^T \prod_{k=\ell}^{2\ell-2l+1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbf{e}_r \mathbb{1}_{\{s=0\}}
\end{aligned}
$$

$$
\begin{aligned}
&+\lim_{n\to\infty}\sum_{l=\lfloor(\ell-1)/2\rfloor+1}^{\ell}\Big(\prod_{j=0}^{\ell-l-2}\hat{\mathrm{B}}^{2l+1,j}\mathrm{B}^{2l+1,j+1}\mathbf{e}_m^T\prod_{k=\ell}^{2\ell-2l}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbf{e}_r\mathbb{1}_{\{s=0\}}\Big)^T\\
&\qquad\big(\mathrm{B}^{2l+1,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{2l+1,-1}\mathrm{B}^{2l+1,0}\\
&\qquad\prod_{j=0}^{\ell-l-2}\hat{\mathrm{B}}^{2l+1,j}\mathrm{B}^{2l+1,j+1}\mathbf{e}_m^T\prod_{k=\ell}^{2\ell-2l}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbf{e}_r\mathbb{1}_{\{s=0\}}\\
&+\sum_{i=1}^{\ell}\prod_{k=\ell}^{i+1}\mathbf{e}_m^T\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\Big(\sum_d d\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}](\mathsf{a}^{(i)}-\mathsf{a}^{(i)}(d))\\
&\qquad+\mathrm{V}^{(i)}\sum_d d\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}](\mathsf{b}^{(i)}-\mathsf{b}^{(i)}(d))\Big)\\
&-\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}\mathbf{e}_m^T\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\big(\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)]\mathsf{a}^{(\ell-j)}-\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}]\big)\\
&-\lim_{n\to\infty}x^{(\ell)}\mathbb{E}[(V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\mu_1^{(\ell)}],
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}]=&x^{(\ell)}\sum_{j=1}^{\ell}\rho'(1)^j\lambda'(1)^{j-1}\lambda_d\\
&+\sum_{j=0}^{\ell}\lambda_d\mathbf{e}_m^T\prod_{k=\ell}^{\ell-j+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}^{(\ell-j)}(d)\\
&+\sum_{j=\ell+1}^{2\ell}\lambda_d(\lambda'(1)\rho'(1))^{j-\ell}\mathbf{e}_m^T\prod_{k=\ell}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}},
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}]=&x^{(\ell)}\sum_{j=0}^{\ell-1}(\rho'(1)\lambda'(1))^j\rho_d\\
&+\sum_{j=1}^{\ell}\rho_d\mathbf{e}_m^T\prod_{k=\ell}^{\ell-j+2}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathrm{V}^{(\ell-j+1)}\mathsf{b}^{(\ell-j+1)}(d)\\
&+\sum_{j=\ell+1}^{2\ell}\rho_d(\lambda'(1)\rho'(1))^{j-\ell-1}\mathbf{e}_m^T\prod_{k=\ell}^{\ell-j+2}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathrm{V}^{(\ell-j+1)}\mathbf{e}_0,
\end{aligned}
$$

$$
\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)]=(\lambda'(1)\rho'(1))^\ell\Big(x^{(\ell)}+\lambda'(1)\mathbf{e}_m^T\prod_{k=\ell}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}\Big),
$$

$$\begin{aligned}&\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}]\\ =&\Big(x^{(\ell)} + \lambda'(1)\mathbf{e}_m^T \prod_{k=\ell}^{1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}\Big)(\lambda'(1)\rho'(1))^j \prod_{k=\ell-j}^{1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}.\end{aligned}$$

The matrices $\mathrm{V}^{(j)}$ and $\mathrm{C}^{(j)}$ are defined in equations (3.9) and (3.8) respectively in Appendix 3.A.1. The matrices $\mathrm{B}^{l,j}$, $\hat{\mathrm{B}}^{l,j}$ and F are defined in equations (3.35), (3.33) and (3.23) respectively in Appendix 3.A.4, and the matrix $\mathrm{P}^{(\ell,\ell-k)}$ is defined in Appendix 4.6. Moreover, the density $\mathsf{a}^{(j)}(d)$ is the outgoing density at a variable node of degree d in iteration j and similarly $\mathsf{b}^{(j)}(d)$ is the outgoing density at a check node of degree d. Finally, $\mathbf{e}_k$ is the unit vector of length $2m+1$ whose k^{th} component is equal to 1.

Example 3.2. *As a first example, we consider the (3,4)-regular ensemble used over a BAWGN channel. The decoding is performed by a QBP decoder with* 15*-quantization levels where the reliability in the decoder is bounded by* 8.12*. In other words, we choose* $\Delta = 1.16$ *and the messages are* $w_i = i\Delta$*, where* $i \in \{-7, \ldots, 7\}$*. The message-passing rules are the QBP message-passing rules defined in Example 3.1. The threshold of this decoder for a BAWGN channel is* $\sigma_{15,8.12}^{BP} \approx 1.2043$*. Figure 3.5 represents the variance* $\mathcal{V}^{(\ell)}$ *as a function of the channel after* $\ell = 0, \ldots, 6$ *iterations. The crosses in Figure 3.5 represent empirical measurements of* $\mathcal{V}^{(\ell)}$ *for* $n = 10^4$*. We see that there is an excellent agreement with the analytic expressions.*

Example 3.3. *In our second example, we consider the irregular ensemble characterized by* $\lambda(x) = 0.06x+0.82x^2+0.12x^3$ *and* $\rho(x) = 0.075x^2+0.8x^3+0.125x^4$*. Again we consider transmission over a BAWGN channel. Assume a quantized BP decoder with* 15*-quantization levels, where the messages are* $w_i = i\Delta$*, with* $\Delta = 1.994$ *and* $i \in \{-7, \ldots, 7\}$*. The corresponding threshold is* $\sigma_{15,13.96}^{BP} \approx 1.0489$*. We plot in Figure 3.6 the variance* $\mathcal{V}^{(\ell)}$ *as a function of the channel after* $\ell = 0, \ldots, 5$ *iterations. We also plot some points which correspond to an empirical evaluation of the variance for* $n = 10^4$*. Again, we see an excellent agreement.*

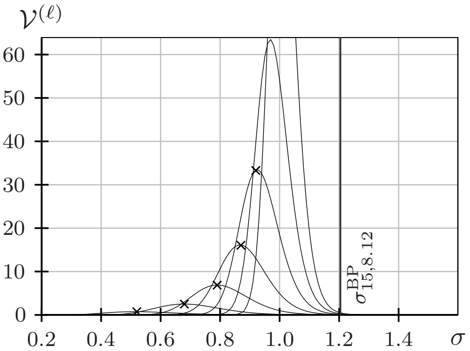

Figure 3.5: The variance $\mathcal{V}^{(\ell)}$ as a function of the channel parameter σ for $\ell = 0, \dots, 6$. The ensemble is $(3, 4)$-regular, transmission takes place over a BAWGN channel, and decoding is accomplished by a quantized BP decoder with 15-quantization levels and a maximum message of ± 8.12. The threshold of this combination is $\sigma^{\mathsf{BP}}_{15,8.12} \approx 1.2043$. The crosses represent the empirically computed variances for $n = 10^4$.

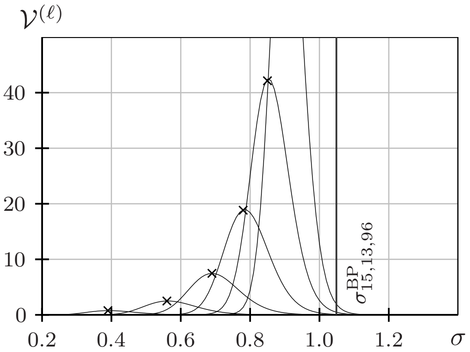

Figure 3.6: The variance $\mathcal{V}^{(\ell)}$ as a function of the channel parameter σ for $\ell = 0, \dots, 5$. The ensemble has degree distribution $\lambda(x) = 0.06x + 0.82x^2 + 0.12x^3$ and $\rho(x) = 0.075x^2 + 0.8x^3 + 0.125x^4$. Transmission takes place over a BAWGN channel and decoding by a quantized BP decoder with 15-quantization levels and a maximum message of ± 13.96. The threshold of this combination is $\sigma^{\mathsf{BP}}_{15,13.96} \approx 1.0489$. The crosses represent the empirically computed variances for $n = 10^4$.

3.A Details of Fixed-ℓ Variance Computations

3.A.1 Edges in $\mathtt{T}_1^{(\ell)}$

$$\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i \in \mathtt{T}_1^{(\ell)}} \mu_1^{(\ell)} \mu_i^{(\ell)} \Big]$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\ell} \mathbb{E}\Big[\sum_{i \in \mathtt{B}_{1,k}} \mu_1^{(\ell)} \mu_i^{(\ell)} \Big]$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbb{E}\Big[\sum_{i \in \mathtt{B}_{1,k}} \mu_1^{(\ell)} \mu_i^{(\ell)} \mid \nu_1^{(\ell-k)} = r \Big] \mathbb{P}\{\nu_1^{(\ell-k)} = r\}$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbb{E}\Big[\sum_{i \in \mathtt{B}_{1,k}} \mu_i^{(\ell)} \mid \nu_1^{(\ell-k)} = r \Big] \mathbb{E}[\mu_1^{(\ell)} \mid \nu_1^{(\ell-k)} = r] \mathbb{P}\{\nu_1^{(\ell-k)} = r\}$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbb{E}\Big[\sum_{i \in \mathtt{B}_{1,k}} \mu_i^{(\ell)} \mid \nu_1^{(\ell-k)} = r \Big] \mathbb{P}\{\nu_1^{(\ell)} = m, \nu_1^{(\ell-k)} = r\}. \quad (3.7)$$

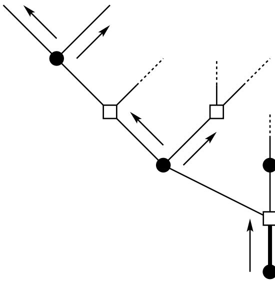

Figure 3.7: Edges in $\mathtt{T}_1^{(2)}$.

We have just seen in Appendix 4.6 how to compute $\mathbb{P}\{\nu_1^{(\ell)}{=}m, \nu_1^{(\ell-k)} = r\}$ (see (4.1)). Let us therefore focus on $\mathbb{E}\Big[\sum_{i\in \mathtt{B}_{1,k}} \mu_i^{(\ell)} \mid \nu_1^{(\ell-k)}{=}r\Big]$.

Assume we are in iteration l. Consider a check node of degree $\mathrm{d}_c{=}d$. Pick one edge connected to this check node and let j be the value of the incoming message on this edge. We want to compute the expected number of outgoing messages (on the $d-1$ *other* connected edges) which are equal to i. This expectation is equal to $(d-1)\mathbb{P}\{\hat{\nu}_{\text{out}}^{(l+1)}{=}i \mid \nu_{\text{in}}^{(l)} = j, \mathrm{d}_c = d\}$. Therefore, if we pick a *random* edge and if we assume that its variable-to-check message has value j, then the expected number of check-to-variable messages of value i on the outgoing edges is equal to $\sum_d \rho_d(d-1)\mathbb{P}\{\hat{\nu}_{\text{out}}^{(l+1)}{=}i \mid \nu_{\text{in}}^{(l)} = j, \mathrm{d}_c = d\}$.

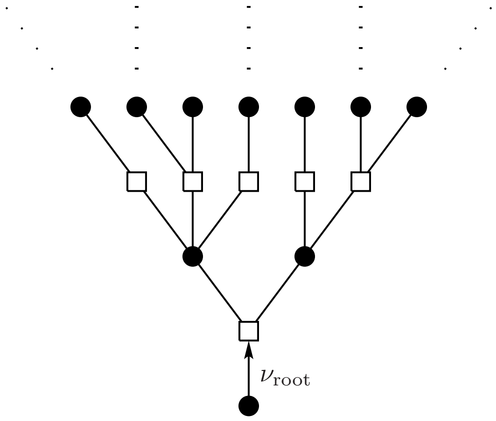

Figure 3.8: Tree which has an edge directed from check to variable as root and whose root node is a variable.

Let us define the $(2m+1)\times(2m+1)$ dimensional matrix $\mathrm{C}^{(l)}$ as

$$\begin{aligned}\mathrm{C}^{(l)}_{i,j} &= \sum_{d} \rho_d (d-1)\mathbb{P}\{\hat{\nu}^{(l+1)}_{\text{out}}{=}i \mid \nu^{(l)}_{\text{in}}{=}j, \mathrm{d}_c{=}d\} \\ &= \sum_{d} \rho_d (d-1) \sum_{j_1,\dots,j_{d\text{-}2}\in\mathcal{W}} \mathbb{1}_{\{\Phi(j,j_1,\dots,j_{d\text{-}2})=i\}} \mathsf{a}^{(l)}_{j_1}\dots\mathsf{a}^{(l)}_{j_{d\text{-}2}}. \end{aligned} \tag{3.8}$$

Define also the vector $\mathsf{c}^{(l)}_k$ of length $2m+1$ whose indices go from $-m$ to m. The i^{th} component of $\mathsf{c}^{(l)}_k$ is the expected number of check-to-variable messages in iteration l which are equal to i, at depth k of a tree which has an edge directed from variable to check node as root and a variable node as root node, as depicted in Figure 3.8. We can then write

$$\mathsf{c}^{(l+1)}_1{=}\mathrm{C}^{(l)}\mathsf{a}^{(l)},$$

where $\mathsf{a}^{(l)}$ is the variable-to-check density in iteration l. In a similar manner, we define the matrix $\mathrm{V}^{(l)}$ for variable nodes. We have

$$\begin{aligned}\mathrm{V}^{(l)}_{i,j} &= \sum_{d} \lambda_d (d-1)\mathbb{P}\{\nu^{(l)}_{\text{out}}{=}i \mid \nu^{(l)}_{\text{in}}{=}j, \mathrm{d}_v{=}d\} \\ &= \sum_{d} \lambda_d (d-1) \sum_{c,j_1,\dots,j_{d\text{-}2}\in\mathcal{W}} \mathbb{1}_{\{\Psi(c,j,j_1,\dots,j_{d\text{-}2})=i\}} (\mathsf{a}_{\mathsf{BMS}})_c\ \mathsf{b}^{(l)}_{j_1}\dots\mathsf{b}^{(l)}_{j_{d\text{-}2}}. \end{aligned} \tag{3.9}$$

Further, let $\mathsf{v}^{(l)}_k$ be the vector whose i^{th} component is the expected number of variable-to-check messages at depth k of the tree which are equal to i in iteration l. We can then write for a full layer

$$\mathsf{v}^{(l+1)}_1 = \mathrm{V}^{(l+1)}\mathrm{C}^{(l)}\mathsf{a}^{(l)}.$$

Let us define $\mathbf{e}_j$ as the unit vector whose j^{th} component is equal to 1. We can then write

$$\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{1,k}} \mu^{(\ell)}_i\Big] = (\mathsf{v}^{(\ell)}_k)_m = \mathbf{e}^T_m \mathsf{v}^{(\ell)}_k = \mathbf{e}^T_m \prod_{j=\ell}^{\ell-k+1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathsf{a}^{(\ell-k)}.$$

If we condition on the value of the messages on the root, we obtain

$$\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{1,k}} \mu^{(\ell)}_i \mid \nu^{(\ell-k)}_1{=}r\Big] = \mathbf{e}^T_m \prod_{j=\ell}^{\ell-k+1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{e}_r.$$

Therefore, we can rewrite (3.7) as

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{T}^{(\ell)}_1} \mu^{(\ell)}_1 \mu^{(\ell)}_i\Big] \\ &= \lim_{n\to\infty} \sum_{k=1}^{\ell} \sum_{r=-m}^{m} \mathbf{e}^T_m \prod_{j=\ell}^{\ell-k+1} \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{e}_r\ \mathbb{P}\{\nu^{(\ell)}_1 = m, \nu^{(\ell-k)}_1 = r\}. \end{aligned} \tag{3.10}$$

3.A.2 Edges in $\mathtt{T}_2^{(\ell)}$

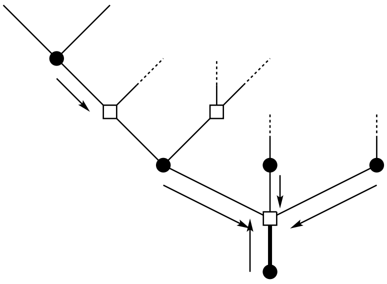

Figure 3.9: Edges in $\mathtt{T}_2^{(2)}$.

Variable-to-check messages carried by edges in $\mathtt{T}_2^{(\ell)}$ in iteration ℓ are independent of the root message in iteration ℓ. Further, the value of the root message as well as the value of messages in $\mathtt{B}_{2,k}$ is independent of the size of this layer. So we have

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_2^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big] &= \lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{2,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ &= \lim_{n\to\infty}(x^{(\ell)})^2\mathbb{E}\Big[|\mathtt{B}_{2,k}|\Big]\\ &= \lim_{n\to\infty}(x^{(\ell)})^2\sum_{k=1}^{\ell}\rho'(1)^k\lambda'(1)^{k-1}. \qquad (3.11)\end{aligned}$$

3.A.3 Edges in $\mathtt{T}_3^{(\ell)}$

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_3^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ &=\lim_{n\to\infty}\sum_{k=1}^{2\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ &=\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]+\lim_{n\to\infty}\sum_{k=\ell+1}^{2\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]. \qquad (3.12)\end{aligned}$$

Let us first look at the first term of the last expression, namely

$$\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big].$$

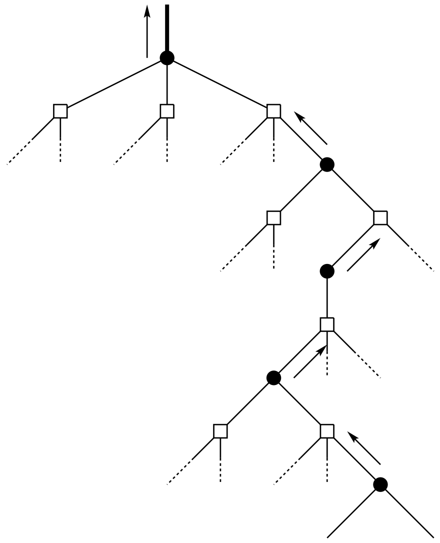

Figure 3.10: Edges in $\mathtt{T}_3^{(2)}$.

We have

$$
\begin{aligned}
&\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
&=\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\mathbb{E}\Big[\mu_1^{(\ell)}\sum_{i\in\mathtt{B}_{3,k}}\mu_i^{(\ell)}\mid|\mathtt{B}_{3,k}|\Big]\Big]\\
&=\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mathbb{E}[\mu_1^{(\ell)}\mu_i^{(\ell)}\mid|\mathtt{B}_{3,k}|\Big]\\
&=\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\sum_{i\in\mathtt{B}_{3,k}}\sum_{r=-m}^{m}\mathbb{E}[\mu_1^{(\ell)}\mu_i^{(\ell)}\mid|\mathtt{B}_{3,k}|,\nu_i^{(\ell-k)}=r]\mathbb{P}\{\nu_i^{(\ell-k)}=r\}\Big]\\
&\overset{\text{(i)}}{=}\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\sum_{i\in\mathtt{B}_{3,k}}\sum_{r=-m}^{m}\mathbb{E}[\mu_1^{(\ell)}\,|\,|\mathtt{B}_{3,k}|,\nu_i^{(\ell-k)}\!=r]\mathbb{E}[\mu_i^{(\ell)}\,|\,\nu_i^{(\ell-k)}\!=r]\mathbb{P}\{\nu_i^{(\ell-k)}\!=r\}\Big]\\
&=\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\sum_{i\in\mathtt{B}_{3,k}}\sum_{r=-m}^{m}\mathbb{E}[\mu_1^{(\ell)}\mid|\mathtt{B}_{3,k}|,\nu_i^{(\ell-k)}=r]\mathbb{P}\{\nu_i^{(\ell)}=m,\nu_i^{(\ell-k)}=r\}\Big]\\
&\overset{\text{(ii)}}{=}\lim_{n\to\infty}\sum_{k=1}^{\ell}\sum_{r=-m}^{m}\mathbb{E}_{|\mathtt{B}_{3,k}|}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mathbb{E}[\mu_1^{(\ell)}\mid|\mathtt{B}_{3,k}|,\nu_i^{(\ell-k)}=r]\Big]\mathbb{P}\{\nu^{(\ell)}=m,\nu^{(\ell-k)}=r\}
\end{aligned}
\tag{3.13}
$$

where in step (i) we have used the fact that knowing $\nu_i^{(\ell-k)}$, where $i \in \mathtt{B}_{3,k}$, $\mu_1^{(\ell)}$ and $\mu_i^{(\ell)}$ become independent. In step (ii) we can drop the index i, since $\mathbb{P}\{\nu_i^{(\ell)} = m, \nu_i^{(\ell-k)} = r\}$ does not depend on i. Let i_1 be the fisrt message in $\mathtt{B}_{3,k}$. Since $\mathbb{E}[\mu_1^{(\ell)} \mid |\mathtt{B}_{3,k}|, \nu_i^{(\ell-k)} = r]$ is invariant for any $i \in \mathtt{B}_{3,k}$, we can then rewrite (3.13) as

$$\begin{aligned}&\lim_{n\to\infty} \sum_{k=1}^{\ell} \mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ &= \lim_{n\to\infty} \sum_{k=1}^{\ell}\sum_{r=-m}^{m} \mathbb{E}[|\mathtt{B}_{3,k}|\mu_1^{(\ell)} \mid \nu_{i_1}^{(\ell-k)} = r]\mathbb{P}\{\nu^{(\ell)}{=}m, \nu^{(\ell-k)}{=}r\} \qquad (3.14)\end{aligned}$$

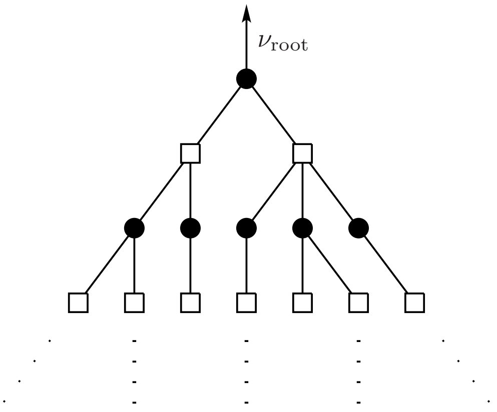

Figure 3.11: Tree which has an edge directed from variable to check as root and whose root node is a variable.

We know how to compute $\mathbb{P}\{\nu^{(\ell)}{=}m, \nu^{(\ell-k)}{=}r\}$ from Appendix 4.6. Further, the derivation of the quantity $\mathbb{E}[|\mathtt{B}_{3,k}|\mu_1^{(\ell)} \mid \nu_{i_1}^{(\ell-k)} = r]$ is similar to the one of $\mathbb{E}[\sum_{i\in\mathtt{B}_{1,k}} \mu_i^{(\ell)} \mid \nu_1^{(\ell-k)} = r]$ in Appendix 3.A.1. We proceed as follows. Consider a variable node of degree d. Pick an edge connected to this variable node and let $\nu_{\text{out}}^{(l)}$ be the variable-to-check message on this edge in iteration l. Moreover let $V = \{\hat{\nu}_{\text{in},1}^{(l)}, \dots, \hat{\nu}_{\text{in},d-1}^{(l)}\}$ be the set of the check-to-variable messages on the $d-1$ remaining edges in teration l (the message indices in the set V are assigned randomly). Let $\hat{\nu}_{\text{in},1}^{(l)} = j$, we want to compute $\mathbb{E}[\mathbb{1}_{\{\nu_{\text{out}}^{(l)}=i\}}|V| \mid \hat{\nu}_{\text{in},1}^{(l)} = j, \mathrm{d}_v = d]$. This is equal to $(d-1)\mathbb{P}\{\nu_{\text{out}}^{(l)} = i \mid \hat{\nu}_{\text{in},1}^{(l)} = j, \mathrm{d}_v = d\}$. Therefore, if we pick a random edge and we look at its message $\nu_{\text{out}}^{(l)}$, we obtain $\mathbb{E}[\mathbb{1}_{\{\nu_{\text{out}}^{(l)}=i\}}|V| \mid \hat{\nu}_{\text{in},1}^{(l)} = j] = \sum_d \lambda_d (d-1)\mathbb{P}\{\nu_{\text{out}}^{(l)} = i \mid \hat{\nu}_{\text{in},1}^{(l)} = j, \mathrm{d}_v = d\}$. Now consider a tree which has an edge directed from variable to check node as root and whose root node is a variable, as depicted in Figure 3.11. Let us define $\mathbf{v}_k^{(l)}$ as the vector whose i^{th} component is equal to the expectation of the number of edges which are directed from variable to check node at depth k of this tree, multiplied by $\mathbb{1}_{\{\nu_{\text{root}}^{(\ell)}=i\}}$, where $\nu_{\text{root}}^{(\ell)}$ is the variable-to-check message on the root. We thus have

$$\mathbf{v}_0^{(\ell)} = \mathrm{V}^{(\ell)}\mathsf{b}^{(\ell)},$$

where $\mathrm{V}_{i,j}^{(\ell)} = \sum_d \lambda_d(d-1)\mathbb{P}\{\nu_{\text{out}}^{(l)} = i \mid \hat{\nu}_{\text{in},1}^{(l)} = j, \mathrm{d}_v = d\}$, as defined in Appendix 3.A.1. In the same way, we also defined the vector $\mathbf{c}_k^{(l)}$ whose i^{th} component is equal to the expectation of the number of edges directed from check to variable node at depth k of the tree, multiplied by $\mathbb{1}_{\{\nu_{\text{root}}^{(\ell)}=i\}}$ as well as a matrix for check nodes, $\mathrm{C}_{i,j}^{(\ell)} = \sum_d \rho_d(d-1)\mathbb{P}\{\hat{\nu}_{\text{out}}^{(l+1)} = i \mid \nu_{\text{in},1}^{(l)} = j, \mathrm{d}_c = d\}$, as defined in Appendix 3.A.1. We can the write for a full layer

$$\mathbf{c}_1^{(\ell)} = \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell-1)}\mathsf{a}^{(\ell-1)},$$

We thus have

$$\mathbb{E}[|\mathsf{B}_{3,k}|\mu_1^{(\ell)}] = (\mathbf{c}_k^{(\ell)})_m = \mathbf{e}_m^T\mathbf{c}_k^{(\ell)} = \mathbf{e}_m^T \prod_{l=\ell}^{\ell-k+1} \mathrm{V}^{(l)}\mathrm{C}^{(l-1)}\mathsf{a}^{(\ell-k)}.$$

Conditionning on $\nu_{i_1}^{(\ell-k)} = r$, we obtain

$$\mathbb{E}[|\mathsf{B}_{3,k}|\mu_1^{(\ell)} \mid \nu_{i_1}^{(\ell-k)} = r] = \mathbf{e}_m^T \prod_{l=\ell}^{\ell-k+1} \mathrm{V}^{(l)}\mathrm{C}^{(l-1)}\mathbf{e}_r. \tag{3.15}$$

So, we can finally write

$$\begin{aligned}&\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathsf{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ &\overset{(3.14,3.15)}{=}\lim_{n\to\infty}\sum_{k=1}^{\ell}\sum_{r=-m}^{m}\mathbf{e}_m^T\prod_{j=\ell}^{\ell-k+1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{e}_r\;\mathbb{P}\{\nu^{(\ell)}=m,\nu^{(\ell-k)}=r\}.\end{aligned} \tag{3.16}$$

Let us then consider the second term of (3.12). More precisely, we consider $\lim_{n\to\infty}\sum_{k=\ell+1}^{2\ell}\mathbb{E}\big[\sum_{i\in\mathsf{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\big]$. Note that $\mu_1^{(\ell)}$ depends on the realization of its computation tree, and thus depends on the size of the boundary of this computation tree, i.e., $\mu_1^{(\ell)}$ depends on $|\mathsf{B}_{3,\ell}|$. Next note that the average number of edges in $\mathsf{B}_{3,k}$, $k > \ell$, is proportional to the average number of edges in $\mathsf{B}_{3,\ell}$. More precisely, $\mathbb{E}[|\mathsf{B}_{3,k}|]=(\rho'(1)\lambda'(1))^{(k-\ell)}\mathbb{E}[|\mathsf{B}_{3,\ell}|]$. Let us define $\mathtt{S}_i^k$ to be the tree of height k whose root is edge i and let $\mathtt{N}(\mathtt{S}_i^k)$ its noise realization. For

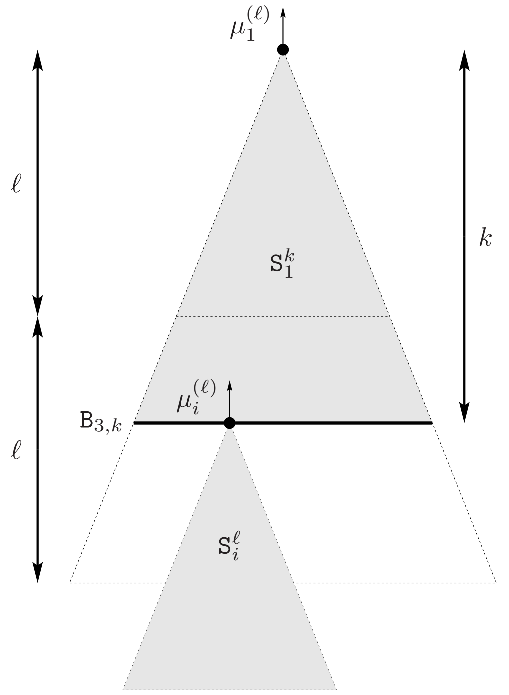

Figure 3.12: $\mathtt{S}_1^k$ and $\mathtt{S}_i^\ell$, with $i \in \mathtt{B}_{3,k}$.

a particular realization $\mathtt{S}_1^k$, we denotes its boundary by $\mathtt{B}_{3,k}(\mathtt{S}_1^k)$. Finally, let

$j_1, \dots, j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}$ be the indices of edges which are in $\mathtt{B}_{3,k}(\mathtt{S}_1^k)$. We have

$$
\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}} \mu_1^{(\ell)}\mu_i^{(\ell)}\Big] \qquad (3.17)\\
&\overset{\text{(i)}}{=} \lim_{n\to\infty} \sum_{\mathtt{S}_1^k} \sum_{\mathtt{N}(\mathtt{S}_1^k)} \sum_{\mathtt{S}_{j_1}^\ell,\dots,\mathtt{S}^\ell_{j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}}} \sum_{\mathtt{N}(\mathtt{S}_{j_1}^\ell),\dots,\mathtt{N}(\mathtt{S}^\ell_{j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}})} \sum_{i\in\mathtt{B}_{3,k}(\mathtt{S}_1^k)}\\
&\qquad \mu_1^{(\ell)}\mu_i^{(\ell)}\mathbb{P}\{\mathtt{S}_1^k\}\mathbb{P}\{\mathtt{N}(\mathtt{S}_1^k)\}\mathbb{P}\{\mathtt{S}_{j_1}^\ell\}\mathbb{P}\{\mathtt{N}(\mathtt{S}_{j_1}^\ell)\}\cdots\mathbb{P}\{\mathtt{S}^\ell_{j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}}\}\mathbb{P}\{\mathtt{N}(\mathtt{S}^\ell_{j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}})\}\\
&\overset{\text{(ii)}}{=} \lim_{n\to\infty} x^{(\ell)} \sum_{\mathtt{S}_1^k} \sum_{\mathtt{N}(\mathtt{S}_1^k)} \sum_{i\in\mathtt{B}_{3,k}(\mathtt{S}_1^k)} \mu_1^{(\ell)}\mathbb{P}\{\mathtt{S}_1^k\}\mathbb{P}\{\mathtt{N}(\mathtt{S}_1^k)\}\\
&= \lim_{n\to\infty} x^{(\ell)} \sum_{\mathtt{S}_1^k} \sum_{\mathtt{N}(\mathtt{S}_1^k)} |\mathtt{B}_{3,k}(\mathtt{S}_1^k)|\mu_1^{(\ell)}\mathbb{P}\{\mathtt{S}_1^k\}\mathbb{P}\{\mathtt{N}(\mathtt{S}_1^k)\}\\
&= \lim_{n\to\infty} x^{(\ell)}\mathbb{E}[|\mathtt{B}_{3,k}|\mu_1^{(\ell)}]\\
&= \lim_{n\to\infty} x^{(\ell)}\mathbb{E}[(\rho'(1)\lambda'(1))^{(k-\ell)}|\mathtt{B}_{3,\ell}|\mu_1^{(\ell)}]\\
&= \lim_{n\to\infty} x^{(\ell)}(\rho'(1)\lambda'(1))^{(k-\ell)}\mathbb{E}[|\mathtt{B}_{3,\ell}|\mu_1^{(\ell)}]\\
&= \lim_{n\to\infty} x^{(\ell)}(\rho'(1)\lambda'(1))^{(k-\ell)}\mathbf{e}_m^T\prod_{j=\ell}^{1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathsf{a}_{\mathsf{BMS}}, \qquad (3.18)
\end{aligned}
$$

where in step (i) we were allowed to write a product of probabilities since $\mathtt{S}_1^k, \mathtt{S}_{j_1}^\ell, \dots, \mathtt{S}^\ell_{j_{|\mathtt{B}_{3,k}(\mathtt{S}_1^k)|}}$ are independent, and in step (ii) we have used the fact that $\sum_{\mathtt{S}^\ell_{j_b}}\sum_{\mathtt{N}(\mathtt{S}^\ell_{j_b})}\mu_i^{(\ell)}\mathbb{P}\{\mathtt{S}^\ell_{j_b}\}\mathbb{P}\{\mathtt{N}(\mathtt{S}^\ell_{j_b})\}=x^{(\ell)}$, for all $b \in \{1, \dots, |\mathtt{B}_{3,k}(\mathtt{S}_1^k)|\}$. The last step is similar to the computation of $\mathbb{E}[|\mathtt{B}_{3,k}|\mu_1^{(\ell)} \mid \nu_{i_1}^{(\ell-k)} = r]$ above.

3.A.4 Edges in $\mathtt{T}_4^{(\ell)}$

In this section, let us use a new notation as depicted in Figure 3.14. In a chain of l layers, we label the variable nodes from 0 to l and the check nodes from $\hat{1}$ to l. Let us also rename the ν's as depicted in Figure 3.14. For instance, $\nu^{(\ell)}_{j\to j\hat{+}1}$ denotes the variable-to-check message in iteration ℓ which is sent from variable node j to check node $j \hat{+} 1$. For a moment let us consider a regular code (d_v,d_c). We also define $\omega_j^{(\ell)}$ and $\omega_{\hat{j}}^{(\ell)}$; $\omega_j^{(\ell)}$ "summarizes" the value received from the channel and the $d_v - 2$ remaining incoming messages at variable node j in iteration ℓ; similarly, $\omega_{\hat{j}}^{(\ell)}$ summarizes the $d_c - 2$ remaining incoming messages at check node $\hat{j}$ in iteration ℓ (see Figure 3.14).

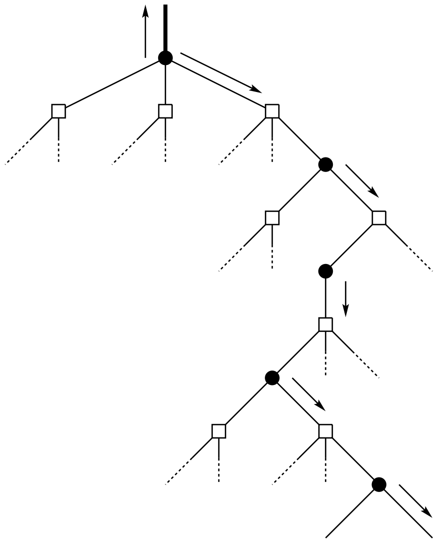

Figure 3.13: Edges in $\mathtt{T}_4^{(2)}$.

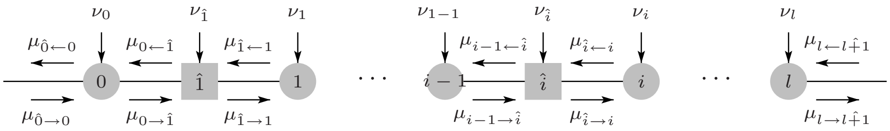

Figure 3.14: Chain of alternating variable and check nodes.

Consider two edges separated by a distance l, where l is even.

$$\begin{aligned}
&\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}=m,\nu^{(\ell)}_{l\to l\hat{+}1}=m\}\\
&=\sum_{r,s}\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}=m,\nu^{(\ell)}_{l\to l\hat{+}1}=m\mid\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\\
&\qquad\mathbb{P}\{\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\\
&=\sum_{r,s}\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}=m\mid\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\mathbb{P}\{\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s\}\\
&\qquad\mathbb{P}\{\nu^{(\ell)}_{l\to l\hat{+}1}=m\mid\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\mathbb{P}\{\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\\
&=\sum_{r,s}\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}=m,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s\mid\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\}\\
&\qquad\mathbb{P}\{\nu^{(\ell)}_{l\to l\hat{+}1}=m,\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\leftarrow l/2}=r\mid\nu^{(\ell\text{-}l/2)}_{l\hat{/}2\to l/2}=s\},
\end{aligned}\tag{3.19}$$

where in the second step we have used the fact that conditioned on $\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2}$ and $\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}$, $\nu^{(\ell)}_{\hat{0}\leftarrow 0}$ and $\nu^{(\ell)}_{l\to \widehat{l+1}}$ are independent and that $\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2}$ and $\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}$ themselves are independent. The computations of the two terms in the last equation are almost identical. Let us consider $\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2} = s \mid \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}{=}r\}$. We have (see Figure 3.15)

$$
\begin{aligned}
&\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2} = s \mid \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2} = r\} \\
&= \sum_{j,k,u,v} \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2}{=}s, \nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\to \widehat{l/2}}{=}j, \nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow \widehat{l/2}}{=}k \mid \\
&\qquad\qquad\qquad \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}{=}r, \omega^{(\ell\text{-}l/2\text{-}1)}_{\widehat{l/2}}{=}u, \omega^{(\ell\text{-}l/2)}_{\widehat{l/2}}{=}v\} \\
&\quad \mathbb{P}\{\omega^{(\ell\text{-}l/2\text{-}1)}_{\widehat{l/2}}{=}u, \omega^{(\ell\text{-}l/2)}_{\widehat{l/2}} v\} \\
&= \sum_{j,k,u,v} \mathbb{P}\{\nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow \widehat{l/2}}{=}k \mid \nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}{=}r, \omega^{(\ell\text{-}l/2)}_{\widehat{l/2}}{=}v\} \\
&\quad \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\to \widehat{l/2}}{=}j \mid \nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow \widehat{l/2}}{=}k\} \\
&\quad \mathbb{P}\{\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2}{=}s \mid \nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\to \widehat{l/2}}{=}j, \omega^{(\ell\text{-}l/2\text{-}1)}_{\widehat{l/2}}{=}u\} \\
&\quad \mathbb{P}\{\omega^{(\ell\text{-}l/2\text{-}1)}_{\widehat{l/2}}{=}u, \omega^{(\ell\text{-}l/2)}_{\widehat{l/2}}{=}v\}. \qquad (3.20)
\end{aligned}
$$

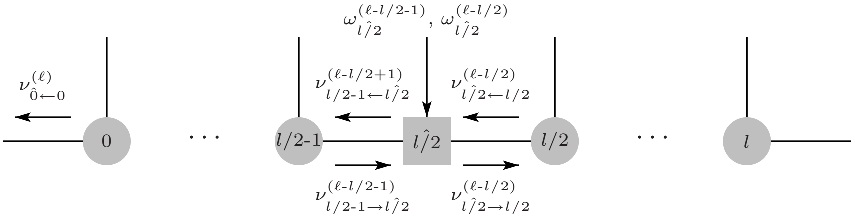
$\omega^{(\ell\text{-}l/2\text{-}1)}_{\widehat{l/2}}$, $\omega^{(\ell\text{-}l/2)}_{\widehat{l/2}}$
$\nu^{(\ell)}_{\hat{0}\leftarrow 0}$
$\nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow \widehat{l/2}}$
$\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\leftarrow l/2}$
0
. . .
$l/2$-1
$\widehat{l/2}$
$l/2$
. . .
l
$\nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\to \widehat{l/2}}$
$\nu^{(\ell\text{-}l/2)}_{\widehat{l/2}\to l/2}$

Figure 3.15: Chain of even length l.

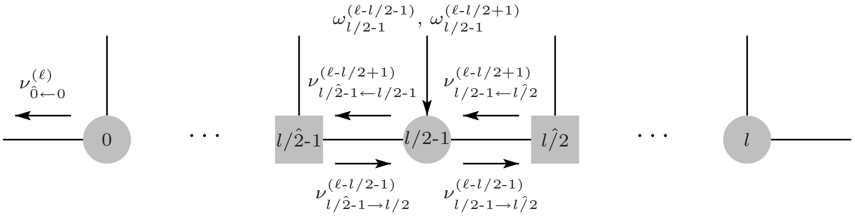
$\omega^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1}$, $\omega^{(\ell\text{-}l/2+1)}_{l/2\text{-}1}$
$\nu^{(\ell)}_{\hat{0}\leftarrow 0}$
$\nu^{(\ell\text{-}l/2+1)}_{l/\hat{2}\text{-}1\leftarrow l/2\text{-}1}$
$\nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow \widehat{l/2}}$
0
. . .
$l/\hat{2}$-1
$l/2$-1
$\widehat{l/2}$
. . .
l
$\nu^{(\ell\text{-}l/2\text{-}1)}_{l/\hat{2}\text{-}1\to l/2}$
$\nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\to \widehat{l/2}}$

Figure 3.16: Chain of even length l.

Similarly (see Figure 3.16),

$$\begin{aligned}
&\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\rightarrow l\hat{/}2}{=}j \mid \nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow l\hat{/}2}{=}k\} \\
={}& \sum_{s,r,u,v} \mathbb{P}\{\nu^{(\ell\text{-}l/2+1)}_{l/\hat{2}\text{-}1\leftarrow l/2\text{-}1}{=}r \mid \nu^{(\ell\text{-}l/2+1)}_{l/2\text{-}1\leftarrow l\hat{/}2}{=}k, \omega^{(\ell\text{-}l/2+1)}_{l/2\text{-}1}{=}v\} \\
&\quad \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}1)}_{l/\hat{2}\text{-}1\rightarrow l/2\text{-}1}{=}s \mid \nu^{(\ell\text{-}l/2+1)}_{l/\hat{2}\text{-}1\leftarrow l/2\text{-}1}{=}r\} \\
&\quad \mathbb{P}\{\nu^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1\rightarrow l\hat{/}2}{=}j \mid \nu^{(\ell\text{-}l/2\text{-}1)}_{l/\hat{2}\text{-}1\rightarrow l/2\text{-}1}{=}s, \omega^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1}{=}u\} \\
&\quad \mathbb{P}\{\omega^{(\ell\text{-}l/2\text{-}1)}_{l/2\text{-}1}{=}u, \omega^{(\ell\text{-}l/2+1)}_{l/2\text{-}1}{=}v\}.
\end{aligned} \tag{3.21}$$

Note that in (3.21), u and v belong to $\{-(d-1)m, \ldots, (d-1)m\}$, according to our definition of the variabe node rule. Let us define the two vectors $\mathbf{c}^{l,i}$ and $\hat{\mathbf{c}}^{l,i}$ of length $(2m+1)^2$, where we denote the index of their y^{th} component by (a, b), such that $y = a(2m+1) + b$, as

$$\begin{aligned}
\mathbf{c}^{l,i}_{(s,r)} &= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}i)}_{l/\hat{2}\text{-}i\rightarrow l/2\text{-}i}{=}s \mid \nu^{(\ell\text{-}l/2+i)}_{l/\hat{2}\text{-}i\leftarrow l/2\text{-}i}{=}r\} \\
\hat{\mathbf{c}}^{l,i}_{(j,k)} &= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}i)}_{l/2\text{-}i\rightarrow l/2\hat{\text{-}}i+1}{=}j \mid \nu^{(\ell\text{-}l/2+i)}_{l/2\text{-}i\leftarrow l/2\hat{\text{-}}i+1}{=}k\}.
\end{aligned}$$

Note that by symmetry we also have

$$\begin{aligned}
\mathbf{c}^{l,i}_{(s,r)} &= \mathbb{P}\{\nu^{(\ell)}_{l\rightarrow l\hat{+}1} = m, \nu^{(\ell\text{-}l/2\text{-}i)}_{l/2+i\leftarrow l/2\hat{+}i+1} = s \mid \nu^{(\ell\text{-}l/2+i)}_{l/2+i\leftarrow l/2\hat{+}i+1}{=}r\} \\
\hat{\mathbf{c}}^{l,i}_{(j,k)} &= \mathbb{P}\{\nu^{(\ell)}_{l\rightarrow l\hat{+}1}{=}m, \nu^{(\ell\text{-}l/2\text{-}i)}_{l/\hat{2}+i\leftarrow l/2+i} = j \mid \nu^{(\ell\text{-}l/2+i)}_{l/\hat{2}+i\rightarrow l/2+i}{=}k\}.
\end{aligned}$$

We also define two $(2m+1)^2 \times (2m+1)^2$ dimensional matrices $\hat{\mathrm{B}}^{l,i}_{(j,k),(r,s)}$ and $\mathrm{B}^{l,i}_{(j,k),(r,s)}$ as

$$\begin{aligned}
\hat{\mathrm{B}}^{l,i}_{(j,k),(r,s)} = \sum_{u,v} &\mathbb{P}\{\nu^{(\ell-l/2-i)}_{l/\hat{2}-i\rightarrow l/2-i} = s \mid \nu^{(\ell-l/2-i-1)}_{l/2-i-1\rightarrow l/\hat{2}-i} = j, \omega^{(\ell-l/2-i-1)}_{l/\hat{2}-i} = u\} \\
&\mathbb{P}\{\nu^{(\ell-l/2+i+1)}_{l/2-i-1\leftarrow l/\hat{2}-i}{=}k \mid \nu^{(\ell-l/2+i)}_{l/\hat{2}-i\leftarrow l/2-i} = r, \omega^{(\ell-l/2+i)}_{l/\hat{2}-i} = v\} \\
&\mathbb{P}\{\omega^{(\ell-l/2-i-1)}_{l/\hat{2}-i} = u, \omega^{(\ell-l/2+i)}_{l/\hat{2}-i} = v\},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{B}^{l,i}_{(j,k),(r,s)} = \sum_{u,v} &\mathbb{P}\{\nu^{(\ell-l/2-i)}_{l/2-i\rightarrow l/2\hat{-}i+1} = j \mid \nu^{(\ell-l/2-i)}_{l/\hat{2}-i\rightarrow l/2-i} = s, \omega^{(\ell-l/2-i)}_{l/2-i} = u\} \\
&\mathbb{P}\{\nu^{(\ell-l/2+i)}_{l/\hat{2}-i\leftarrow l/2-i}{=}r \mid \nu^{(\ell-l/2+i)}_{l/2-i\leftarrow l/2\hat{-}i+1} = k, \omega^{(\ell-l/2+i)}_{l/2-i} = v\} \\
&\mathbb{P}\{\omega^{(\ell-l/2-i)}_{l/2-i} = u, \omega^{(\ell-l/2+i)}_{l/2-i} = v\}.
\end{aligned}$$

According to (3.20) and (3.21), we can write $\mathbf{c}^{l,i}{=}\hat{\mathrm{B}}^{l,i}\hat{\mathbf{c}}^{l,i+1} = \hat{\mathrm{B}}^{l,i}\mathrm{B}^{l,i+1}\mathbf{c}^{l,i+1}$. Thus we can expand (3.19) as follows

$$\begin{aligned}
&\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell)}_{l\rightarrow \widehat{l+1}} = m\}\\
\overset{(3.19)}{=}&\sum_{r,s}\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-l/2)}_{\widehat{l/2}\rightarrow l/2} = s \mid \nu^{(\ell-l/2)}_{\widehat{l/2}\leftarrow l/2} = r\}\\
&\quad\mathbb{P}\{\nu^{(\ell)}_{l\rightarrow \widehat{l+1}}{=}m, \nu^{(\ell-l/2)}_{\widehat{l/2}\leftarrow l/2} = r \mid \nu^{(\ell-l/2)}_{\widehat{l/2}\rightarrow l/2} = s\}\\
=&\sum_{r,s}\mathbf{c}^{l,0}_{(s,r)}\hat{\mathbf{c}}^{l,0}_{(r,s)}\\
=&(\mathbf{c}^{l,0})^T\mathrm{F}\hat{\mathbf{c}}^{l,0}\\
=&(\mathbf{c}^{l,0})^T\mathrm{F}\mathrm{B}^{l,0}\mathbf{c}^{l,0}\\
=&\big(\hat{\mathrm{B}}^{l,0}\mathrm{B}^{l,1}\cdots\hat{\mathrm{B}}^{l,\min\{l/2,\ell-l/2\}-1}\mathrm{B}^{l,\min\{l/2,\ell-l/2\}}\mathbf{c}^{l,\min\{l/2,\ell-l/2\}}\big)^T\\
&\mathrm{F}\mathrm{B}^{l,0}\hat{\mathrm{B}}^{l,0}\mathrm{B}^{l,1}\cdots\hat{\mathrm{B}}^{l,\min\{l/2,\ell-l/2\}-1}\mathrm{B}^{l,\min\{l/2,\ell-l/2\}}\mathbf{c}^{l,\min\{l/2,\ell-l/2\}},\\
=&\Big(\prod_{j=0}^{\min\{l/2,\ell-l/2\}-1}\hat{\mathrm{B}}^{l,j}\mathrm{B}^{l,j+1}\mathbf{c}^{l,\min\{l/2,\ell-l/2\}}\Big)^T\\
&\mathrm{F}\mathrm{B}^{l,0}\prod_{j=0}^{\min\{l/2,\ell-l/2\}-1}\hat{\mathrm{B}}^{l,j}\mathrm{B}^{l,j+1}\mathbf{c}^{l,\min\{l/2,\ell-l/2\}},
\end{aligned}\tag{3.22}$$

where F is a $(2m+1)^2 \times (2m+1)^2$ permutation matrix which switches the order of the two indices. More precisely,

$$\mathrm{F}_{(i,j),(k,l)} = \begin{cases} 1 & \text{if } i = l \text{ and } j = k,\\ 0 & \text{otherwise.}\end{cases}\tag{3.23}$$

If $l \leq \ell$:

$$\begin{aligned}
\mathbf{c}^{l,min\{l/2,\ell-l/2\}}_{(s,r)} &= \mathbf{c}^{l,l/2}_{(s,r)}\\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-l)}_{\hat{0}\rightarrow 0} = s \mid \nu^{(\ell)}_{\hat{0}\leftarrow 0} = r\}\\
&= \mathbb{P}\{\nu^{(\ell-l)}_{\hat{0}\rightarrow 0} = s\}\mathbb{1}_{\{m=r\}}\\
&= \mathbf{u}_s^T\mathsf{b}^{\ell-l}\mathbb{1}_{\{m=r\}}.
\end{aligned}\tag{3.24}$$

If $l > \ell$:

$$\begin{aligned}
\mathbf{c}^{l,min\{l/2,\ell-l/2\}}_{(s,r)} &= \mathbf{c}^{l,\ell-l/2}_{(s,r)}\\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(0)}_{\widehat{l-\ell}\rightarrow l-\ell} = s \mid \nu^{(2\ell-l)}_{\widehat{l-\ell}\leftarrow l-\ell} = r\}\\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m \mid \nu^{(2\ell-l)}_{\widehat{l-\ell}\leftarrow l-\ell} = r\}\mathbb{1}_{\{s=0\}}\\
&= \mathbf{u}_m^T\prod_{j=\ell}^{2\ell-l+1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{u}_r\mathbb{1}_{\{s=0\}}.
\end{aligned}\tag{3.25}$$

So far we have considered the case of even l. If l is odd, we define

$$c^{l,i}_{(s,r)} = \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-(l+1)/2-i)}_{(l-1\hat{)}/2-i\rightarrow(l-1)/2-i} = s \mid \nu^{(\ell-(l-1)/2+i)}_{(l-1\hat{)}/2-i\leftarrow(l-1)/2-i} = r\}$$
$$\hat{c}^{l,i}_{(j,k)} = \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-(l+1)/2-i)}_{(l-1)/2-i\rightarrow(l-1)\hat{/}2-i+1} = j \mid \nu^{(\ell-(l-1)/2+i)}_{(l-1)/2-i\leftarrow(l-1)\hat{/}2-i+1} = k\},$$

and

$$\begin{aligned}
&\hat{\mathrm{B}}^{l,i}_{(j,k),(r,s)}\\
&= \sum_{u,v} \mathbb{P}\{\nu^{(\ell-(l+1)/2-i)}_{(l-1\hat{)}/2-i\rightarrow(l-1)/2-i} = s \mid \nu^{(\ell-(l+1)/2-i-1)}_{(l-1)/2-i-1\rightarrow(l-1\hat{)}/2-i} = j, \omega^{(\ell-(l+1)/2-i-1)}_{(l-1\hat{)}/2-i} = u\}\\
&\qquad \mathbb{P}\{\nu^{(\ell-(l-1)/2+i+1)}_{(l-1)/2-i-1\leftarrow(l-1\hat{)}/2-i}{=}k \mid \nu^{(\ell-(l-1)/2+i)}_{(l-1\hat{)}/2-i\leftarrow(l-1)/2-i} = r, \omega^{(\ell-(l-1)/2+i)}_{(l-1\hat{)}/2-i} = v\}\\
&\qquad \mathbb{P}\{\omega^{(\ell-(l+1)/2-i-1)}_{(l-1\hat{)}/2-i}{=}u, \omega^{(\ell-(l-1)/2+i)}_{(l-1\hat{)}/2-i} = v\}\\
&\mathrm{B}^{l,i}_{(j,k),(r,s)}\\
&= \sum_{u,v} \mathbb{P}\{\nu^{(\ell-(l+1)/2-i)}_{(l-1)/2-i\rightarrow(l-1)\hat{/}2-i+1} = j \mid \nu^{(\ell-(l+1)/2-i)}_{(l-1\hat{)}/2-i\rightarrow(l-1)/2-i} = s, \omega^{(\ell-(l+1)/2-i)}_{(l-1)/2-i} = u\}\\
&\qquad \mathbb{P}\{\nu^{(\ell-(l-1)/2+i)}_{(l-1\hat{)}/2-i\leftarrow(l-1)/2-i}{=}r \mid \nu^{(\ell-(l-1)/2+i)}_{(l-1)/2-i\leftarrow(l-1)\hat{/}2-i+1} = k, \omega^{(\ell-(l-1)/2+i)}_{(l-1)/2-i} = v\}\\
&\qquad \mathbb{P}\{\omega^{(\ell-(l+1)/2-i)}_{(l-1)/2-i}{=}u, \omega^{(\ell-(l-1)/2+i)}_{(l-1)/2-i} = v\}.
\end{aligned}$$

By symmetry we have

$$c^{l,i}_{(s,r)} = \mathbb{P}\{\nu^{(\ell)}_{l\rightarrow l\hat{+}1} = m, \nu^{(\ell-(l+1)/2-i)}_{(l+1)/2+i\leftarrow(l+1)\hat{/}2+i+1} = s \mid \nu^{(\ell-(l-1)/2+i)}_{(l+1)/2+i\rightarrow(l+1\hat{)}/2+i} = r\}$$
$$\hat{c}^{l,i}_{(j,k)} = \mathbb{P}\{\nu^{(\ell)}_{l\leftarrow l\hat{+}1} = m, \nu^{(\ell-(l+1)/2-i)}_{(l+1\hat{)}/2+i\leftarrow(l+1)/2+i} = j \mid \nu^{(\ell-(l-1)/2+i)}_{(l+1\hat{)}/2+i\rightarrow(l+1)/2+i} = k\}.$$

Then

$$\begin{aligned}
&\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell)}_{l\to\widehat{l+1}} = m\} \\
={}&\sum_{r,s}\mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-(l+1)/2)}_{(l-1)/2\to\widehat{(l-1)/2+1}} = s \mid \nu^{(\ell-(l-1)/2)}_{(l-1)/2\leftarrow\widehat{(l-1)/2+1}} = r\} \\
&\quad\mathbb{P}\{\nu^{(\ell)}_{l\to\widehat{l+1}}{=}m, \nu^{(\ell-(l-1)/2)}_{\widehat{(l-1)/2}\leftarrow(l-1)/2+1} = r \mid \nu^{(\ell-(l+1)/2)}_{\widehat{(l-1)/2}\to(l-1)/2} = s\} \\
={}&\sum_{r,s}\hat{\mathbf{c}}^{l,0}_{(s,r)}\mathbf{c}^{l,-1}_{(r,s)} \\
={}&(\hat{\mathbf{c}}^{l,0})^T\mathrm{F}\mathbf{c}^{l,-1} \\
={}&\big(\mathrm{B}^{l,0}\mathbf{c}^{l,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{l,-1}\mathrm{B}^{l,0}\mathbf{c}^{l,0} \\
={}&(\mathbf{c}^{l,0})^T\big(\mathrm{B}^{l,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{l,-1}\mathrm{B}^{l,0}\mathbf{c}^{l,0} \\
={}&\Big(\prod_{j=0}^{\min\{(l-1)/2,\ell-(l+1)/2\}-1}\hat{\mathrm{B}}^{l,j}\mathrm{B}^{l,j+1}\mathbf{c}^{l,\min\{(l-1)/2,\ell-(l+1)/2\}}\Big)^T \\
&\quad\big(\mathrm{B}^{l,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{l,-1}\mathrm{B}^{l,0}\prod_{j=0}^{\min\{(l-1)/2,\ell-(l+1)/2\}-1}\hat{\mathrm{B}}^{l,j}\mathrm{B}^{l,j+1}\mathbf{c}^{l,\min\{(l-1)/2,\ell-(l+1)/2\}},
\end{aligned} \tag{3.26}$$

If $l \leq \ell$:

$$\begin{aligned}
\mathbf{c}^{l,min\{(l-1)/2,\ell-(l+1)/2\}}_{(s,r)} &= \mathbf{c}^{l,(l-1)/2}_{(s,r)} \\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(\ell-l)}_{\hat{0}\to 0} = s \mid \nu^{(\ell)}_{\hat{0}\leftarrow 0} = r\} \\
&= \mathbb{P}\{\nu^{(\ell-l)}_{\hat{0}\to 0} = s\}\mathbb{1}_{\{m=r\}} \\
&= \mathbf{u}_s^T\mathsf{b}^{\ell-l}\mathbb{1}_{\{m=r\}}.
\end{aligned} \tag{3.27}$$

If $l > \ell$:

$$\begin{aligned}
\mathbf{c}^{l,min\{(l-1)/2,\ell-(l+1)/2\}}_{(s,r)} &= \mathbf{c}^{l,\ell-(l+1)/2}_{(s,r)} \\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m, \nu^{(0)}_{\widehat{l-\ell}\to l-\ell} = s \mid \nu^{(2\ell-l)}_{\widehat{l-\ell}\leftarrow l-\ell} = r\} \\
&= \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0} = m \mid \nu^{(2\ell-l)}_{\widehat{l-\ell}\leftarrow l-\ell} = r\}\mathbb{1}_{\{s=0\}} \\
&= \mathbf{u}_m^T\prod_{j=\ell}^{2\ell-l+1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{u}_r\mathbb{1}_{\{s=0\}}.
\end{aligned} \tag{3.28}$$

In order to simplify the notation, let us redefine $\mathbf{c}^{l,i}$ such that it holds for even and odd l:

If $l \leq \ell$:

$$\mathbf{c}^{l,i}_{(s,r)} = \mathbb{P}\{\nu^{(\ell)}_{\hat{0}\leftarrow 0}{=}m, \nu^{(\ell\text{-}l/2\text{-}i)}_{\widehat{l/2\text{-}i}\to l/2\text{-}i}{=}s \mid \nu^{(\ell\text{-}l/2+i)}_{\widehat{l/2\text{-}i}\leftarrow l/2\text{-}i}{=}r\} \tag{3.29}$$

$$= \mathbf{u}_s^T\mathsf{b}^{\ell-l}\mathbb{1}_{\{m=r\}}. \tag{3.30}$$

If $l > \ell$:

$$\mathbf{c}^{l,i} = \mathbf{u}_m^T \prod_{j=\ell}^{2\ell-l+1} \mathrm{V}^{(j)} \mathrm{C}^{(j-1)} \mathbf{u}_r \mathbb{1}_{\{s=0\}}. \tag{3.31}$$

So far we have consider two variable nodes separated by a distance l for a regular code (d_v, d_c). But in order to expand $\lim_{n\to\infty} \mathbb{E}\big[\sum_{i\in \mathtt{T}_4^{(\ell)}} \mu_1^{(\ell)} \mu_i^{(\ell)}\big]$, we need to know how to compute the correlation between the root messages and all messages in $\mathtt{B}_{4,l}$, i.e., $\lim_{n\to\infty} \mathbb{E}\big[\sum_{i\in \mathtt{B}_{4,l}} \mu_1^{(\ell)} \mu_i^{(\ell)}\big]$. This derivation is similar to the one above when we considered only the correlation between two messages, i.e., $\mathbb{P}\{\nu_{\hat{0}\leftarrow 0}^{(\ell)} = m, \nu_{l\rightarrow \hat{l+1}}^{(\ell)} = m\}$, except that we need to add a factor $d_v - 1$ in B and $d_c - 1$ in $\hat{\mathrm{B}}$ in order to take into account the number of messages in $\mathtt{B}_{4,l}$. Finally for irregular ensembles, we need to average the degrees. Thus we redefine the two matrices

$$\begin{aligned}
&\hat{\mathrm{B}}^{l,i}_{(j,k),(r,s)} \qquad\qquad (3.32)\\
&= \sum_d (d-1)\rho_d \\
&\quad \sum_{u,v} \mathbb{P}\{\nu^{(\ell-\lfloor (l+1)/2\rfloor - i)}_{\widehat{\lceil (l-1)/2\rceil} - i \rightarrow \lceil (l-1)/2\rceil - i} = s \,|\, \nu^{(\ell-\lfloor (l+1)/2\rfloor - i-1)}_{\lceil (l-1)/2\rceil - i - 1 \rightarrow \widehat{\lceil (l-1)/2\rceil} - i} = j, \omega^{(\ell-\lfloor (l+1)/2\rfloor - i-1)}_{\widehat{\lceil (l-1)/2\rceil} - i} = u\} \\
&\qquad \mathbb{P}\{\nu^{(\ell-\lceil (l-1)/2\rceil + i+1)}_{\lceil (l-1)/2\rceil - i - 1 \leftarrow \widehat{\lceil (l-1)/2\rceil} - i} = k \,|\, \nu^{(\ell-\lceil (l-1)/2\rceil + i)}_{\widehat{\lceil (l-1)/2\rceil} - i \leftarrow \lceil (l-1)/2\rceil - i} = r, \omega^{(\ell-\lceil (l-1)/2\rceil + i)}_{\widehat{\lceil (l-1)/2\rceil} - i} = v\} \\
&\qquad \mathbb{P}\{\omega^{(\ell-\lfloor (l+1)/2\rfloor - i-1)}_{\widehat{\lceil (l-1)/2\rceil} - i} = u, \omega^{(\ell-\lceil (l-1)/2\rceil + i)}_{\widehat{\lceil (l-1)/2\rceil} - i} = v \,|\, \mathrm{d}_c = d\} \\
&\overset{(i)}{=} \sum_d (d-1)\rho_d \sum_{u,v} \mathbb{1}_{\{s=j\boxplus u\}} \mathbb{1}_{\{k=r\boxplus v\}} \\
&\qquad \big(\mathrm{P}^{(\ell-\lfloor (l+1)/2\rfloor - i-1, \ell-\lceil (l-1)/2\rceil + i)}\big)^{\boxdot (d\text{-}2)}\big)_{u,v}, \qquad\qquad (3.33)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{B}^{l,i}_{(j,k),(r,s)} \qquad\qquad (3.34)\\
&= \sum_d (d-1)\lambda_d \\
&\quad \sum_{u,v} \mathbb{P}\{\nu^{(\ell-\lfloor (l+1)/2\rfloor - i)}_{\lceil (l-1)/2\rceil - i \rightarrow \widehat{\lceil (l-1)/2\rceil} - i+1} = j \,|\, \nu^{(\ell-\lfloor (l+1)/2\rfloor - i)}_{\widehat{\lceil (l-1)/2\rceil} - i \rightarrow \lceil (l-1)/2\rceil - i} = s, \omega^{(\ell-\lfloor (l+1)/2\rfloor - i)}_{\lceil (l-1)/2\rceil - i} = u\} \\
&\qquad \mathbb{P}\{\nu^{(\ell-\lceil (l-1)/2\rceil + i)}_{\widehat{\lceil (l-1)/2\rceil} - i \leftarrow \lceil (l-1)/2\rceil - i} = r \,|\, \nu^{(\ell-\lceil (l-1)/2\rceil + i)}_{\lceil (l-1)/2\rceil - i \leftarrow \widehat{\lceil (l-1)/2\rceil} - i+1} = k, \omega^{(\ell-\lceil (l-1)/2\rceil + i)}_{\lceil (l-1)/2\rceil - i} = v\} \\
&\qquad \mathbb{P}\{\omega^{(\ell-\lfloor (l+1)/2\rfloor - i)}_{\lceil (l-1)/2\rceil - i} = u, \omega^{(\ell-\lceil (l-1)/2\rceil + i)}_{\lceil (l-1)/2\rceil - i} = v \mid \mathrm{d}_v = d\} \\
&\overset{(ii)}{=} \sum_d (d-1)\lambda_d \sum_{u,v} \mathbb{1}_{\{j=\mathfrak{C}(s+u)\}} \mathbb{1}_{\{r=\mathfrak{C}(k+v)\}} \\
&\qquad \Big(\mathfrak{C}\big(\mathrm{C}_{\mathsf{BMS}} \odot (\mathrm{Q}^{(\ell-\lfloor (l+1)/2\rfloor - i, \ell-\lceil (l-1)/2\rceil + i)})^{\odot (d\text{-}2)}\big)\Big)_{u,v}, \qquad\qquad (3.35)
\end{aligned}$$

where steps (i) and (ii) are performed according to Appendix 4.6. Note that in order to simplify the notation we have defined the matrices above such that they hold for even and odd l. Then we have

$$
\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_4^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
&=\lim_{n\to\infty}\mathbb{E}\Big[\sum_{l=0}^{2\ell}\sum_{i\in\mathtt{B}_{4,l}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
&=\lim_{n\to\infty}\Bigg(\left.\begin{aligned}&\sum_{h=0}^{\lceil(\ell\text{-}1)/2\rceil}\Big(\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,h}\Big)^T\\
&\big(\mathrm{F}\mathrm{B}^{2h,0}\big)\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,h}\end{aligned}\right\}\begin{array}{c}\text{length } l=2h\le\ell\\(3.22,3.24)\end{array}\\
&\quad+\left.\begin{aligned}&\sum_{h=0}^{\lfloor(\ell\text{-}1)/2\rfloor}\Big(\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,h}\Big)^T\\
&\big(\mathrm{B}^{2h+1,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{2h+1,\text{-}1}\mathrm{B}^{2h+1,0}\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,h}\end{aligned}\right\}\begin{array}{c}\text{length } l=2h+1\le\ell\\(3.26,3.27)\end{array}\\
&\quad+\left.\begin{aligned}&\sum_{h=\lceil(\ell\text{-}1)/2\rceil+1}^{\ell}\Big(\prod_{j=0}^{\ell\text{-}h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,\ell-h}\Big)^T\\
&\mathrm{F}\mathrm{B}^{2h,0}\prod_{j=0}^{\ell\text{-}h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,\ell-h}\end{aligned}\right\}\begin{array}{c}\text{length } l=2h>\ell\\(3.22,3.25)\end{array}\\
&\quad+\left.\begin{aligned}&\sum_{h=\lfloor(\ell\text{-}1)/2\rfloor+1}^{\ell}\Big(\prod_{j=0}^{\ell\text{-}h\text{-}2}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,\ell-(h+1)}\Big)^T\\
&\big(\mathrm{B}^{2h+1,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{2h+1,\text{-}1}\mathrm{B}^{2h+1,0}\\
&\prod_{j=0}^{\ell\text{-}h\text{-}2}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,\ell-(h+1)}\Bigg).\end{aligned}\right\}\begin{array}{c}\text{length } l=2h+1>\ell\\(3.26,3.28)\end{array}
\end{aligned}
$$

3.A.5 Computation of S^c

Let us now concentrate on the sixth term

$$S^c=\lim_{n\to\infty}\Bigg(\mathbb{E}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]-n\Lambda'(1)(x^{(\ell)})^2\Bigg).$$

At first one might think that the root message and a message on an edge in $(\mathtt{T}^{(\ell)})^c$ are uncorrelated since their computation graphs do not intersect. Indeed, this is the case for a regular ensemble, for which S^c is equal to $-\mid\mathtt{T}^{(\ell)}\mid(x^{(\ell)})^2$. The computation of S^c for irregular ensembles is more challenging. The number of edges in $\mathtt{T}^{(\ell)}$ for irregular ensembles is not constant but is a random variable which depends on the graph realization. Thus we cannot move the sum over $i \in (\mathtt{T}^{(\ell)})^c$ outside the expectation. It is clear that $\mu_1^{(\ell)}$ depends on $\mathrm{G}_{\mathtt{T}}^{(\ell)}$, but as we will see in the two following sections, $\mu_i^{(\ell)}$ also depends on $\mathrm{G}_{\mathtt{T}}^{(\ell)}$. It is therefore natural to condition on $\mathrm{G}_{\mathtt{T}}^{(\ell)}$. We have

$$\begin{aligned}
S^c &= \lim_{n\to\infty}\Big(\mathbb{E}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big)\\
&= \lim_{n\to\infty}\Big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\Big[\,\mathbb{E}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c}\mu_1^{(\ell)}\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\Big]\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big)\\
&= \lim_{n\to\infty}\Big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\Big[\sum_{i\in(\mathtt{T}^{(\ell)})^c}\mathbb{E}\Big[\,\mu_1^{(\ell)}\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\Big]\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big)\\
&= \lim_{n\to\infty}\Big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\Big[\,\mathbb{E}\Big[\,\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\Big]\sum_{i\in(\mathtt{T}^{(\ell)})^c}\mathbb{E}\Big[\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\Big]\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big).
\end{aligned} \tag{3.36}$$

Degree Distribution Correction

The message of the root node is a function of the specific realization of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$. For the regular case, if we consider a fixed number of iterations and the limit of n tending to infinity, there is only one $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ which has a positive probability (namely a regular tree of the appropriate height). But for the irregular case many such computation graphs have a strictly positive probability in the asymptotic limit.

Suppose, e.g., that $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ contains an unusual large number of variable nodes of high degree (as compared to $\lambda(x)$). In this case we expect the average (over the noise realization) reliability of the message emitted by the root node to be higher than what is predicted by density evolution.

But $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ indirectly also influences the messages in $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$. This is true since the total number of nodes of a given degree is fixed. Therefore, in the above case we know that $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ contains fewer variable nodes of high degree than expected. This causes a small deviation of the degree distribution of $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ as compared to $\lambda(x)$ and, hence, a small deviation of average message in $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ as compared to density evolution. Even though this deviation is only of order $1/n$, there are of order n messages inside $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ and so this deviation gives a non-vanishing contribution in the limit of infinite blocklengths. pp Let us write down this effect explicitly. Given the degree distribution polynomials $\lambda(x)$ and $\rho(x)$, define the operators $\lambda(\mathsf{x}) = \sum_i \lambda_i \mathsf{x}^{\star(i-1)}$ and $\rho(\mathsf{x}) = \sum_i \rho_i \mathsf{x}^{\boxast(i-1)}$, where x is a density, $\star$ is the convolution at variable nodes, and $\boxast$ is the convolution at check nodes. This extends to the respective derivatives in the natural way: $\lambda'(\mathsf{x}) = \sum_i (i-1)\lambda_i \mathsf{x}^{\star(i-2)}$ and $\rho'(\mathsf{x}) = \sum_i (i-1)\rho_i \mathsf{x}^{\boxast(i-2)}$. Let us also define the two operators $V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{x}) = \sum_i V_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}} \mathsf{x}^{\star i}$ and $C^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{x}) = \sum_i C_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}} \mathsf{x}^{\boxast i}$, where $V_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}$ and $C_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}$ are the number of variable nodes and check nodes of degree

i in $\mathrm{G}_{\mathtt{T}}^{(\ell)}$, respectively. Again let us extend the notation to their respective derivatives $(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(\mathsf{x}) = \sum_i iV_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}} \mathsf{x}^{\star(i-1)}$ and $(C^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(\mathsf{x}) = \sum_i iC_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}} \mathsf{x}^{\boxast(i-1)}$.

Consider the degree distribution of $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$. We know the overall degree distribution and we are given the degree distribution of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ itself. By removing $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ from the overall bipartite graph the distribution of $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ changes by $\Delta\lambda(x)$ and $\Delta\rho(x)$, respectively. Let us compute these deviations. Consider a bipartite graph of variable and check degree distribution $\lambda(x)$ and $\rho(\mathsf{x})$, respectively. Assume that we remove a variable node of degree j from this graph. Thus the total number of edges becomes $n\Lambda'(1) - j$ and the new variable degree distribution of the graph, call it $\tilde{\lambda}(x)$, is

$$\begin{aligned}
\tilde{\lambda}(x) &= \sum_i \tilde{\lambda}_i x^{i-1} \\
&= \sum_i \frac{n\Lambda'(1)\lambda_i x^{i-1} - jx^{j-1}}{n\Lambda'(1) - j} \\
&= \sum_i \frac{(n\Lambda'(1) - j)\lambda_i x^{i-1} + j\lambda_i x^{i-1} - jx^{j-1}}{n\Lambda'(1) - j} \\
&= \lambda(x) + \frac{j\lambda(x) - jx^{j-1}}{n\Lambda'(1) - j} \\
&= \lambda(x) + \frac{j\lambda(x) - jx^{j-1}}{n\Lambda'(1)} + O(1/n^2).
\end{aligned}$$

The number of edges in $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ is equal to $(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1) = \sum_i iV_i^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}$. Therefore, if we remove $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ from the complete graph the variable degree distribution changes by

$$\Delta\lambda(x) = \frac{(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\lambda(x) - (V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(x)}{n\Lambda'(1)} + O(1/n^2). \tag{3.37}$$

Similarly, the check degree distribution changes by

$$\Delta\rho(x) = \frac{(C^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\rho(x) - (C^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(x)}{n\Lambda'(1)} + O(1/n^2). \tag{3.38}$$

Let us write the variable and the check degree distribution of $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ as $\tilde{\lambda}(x) = \lambda(x) + \Delta\lambda(x)$ and $\tilde{\rho}(x) = \rho(x) + \Delta\rho(x)$, respectively.

Consider the effect of this small deviation on density evolution. As defined earlier, let $\mathsf{a}^{(j)}$ be the variable-to-check node message density in iteration j and let $\tilde{\Delta}\mathsf{a}^{(j)}$ be the deviation of this quantity due to the deviation of the degree distribution. Let $\mathsf{b}^{(j)}$ and $\tilde{\Delta}\mathsf{b}^{(j)}$ be the equivalent quantities for messages flowing from check to variable nodes. Assume that $\tilde{\Delta}\mathsf{a}^{(j)}$ and $\tilde{\Delta}\mathsf{b}^{(j)}$ are of order $O(1/n)$ (the recursion and the initialization will show that this is indeed the case). Consider first the evolution at check nodes. We have

$$\begin{aligned}
&\mathsf{b}^{(j)} - \tilde{\Delta}\mathsf{b}^{(j)}\\
=&\tilde{\rho}(\mathsf{a}^{(j-1)} - \tilde{\Delta}\mathsf{a}^{(j-1)})\\
=&\sum_i \tilde{\rho}_i(\mathsf{a}^{(j-1)} - \tilde{\Delta}\mathsf{a}^{(j-1)})^{\boxast(i-1)}\\
=&\sum_i \tilde{\rho}_i\big((\mathsf{a}^{(j-1)})^{\boxast(i-1)} - (i-1)\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast (\mathsf{a}^{(j-1)})^{\boxast(i-2)} + O(1/n^2)\big)\\
=&\sum_i \tilde{\rho}_i(\mathsf{a}^{(j-1)})^{\boxast(i-1)} - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \sum_i \tilde{\rho}_i(i-1)(\mathsf{a}^{(j-1)})^{\boxast(i-2)} + O(1/n^2)\\
=&\tilde{\rho}(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}) + O(1/n^2).
\end{aligned} \tag{3.39}$$

Next comes the evolution of the densities at the varible-node side. We have

$$\begin{aligned}
&\mathsf{a}^{(j)} - \tilde{\Delta}\mathsf{a}^{(j)}\\
=&\mathsf{a}_{\mathsf{BMS}} \star \tilde{\lambda}(\mathsf{b}^{(j)} - \tilde{\Delta}\mathsf{b}^{(j)})\\
\overset{(3.39)}{=}&\mathsf{a}_{\mathsf{BMS}} \star \tilde{\lambda}(\tilde{\rho}(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}) + O(1/n^2))\\
=&\mathsf{a}_{\mathsf{BMS}} \star \tilde{\lambda}(\rho(\mathsf{a}^{(j-1)}) + \Delta\rho(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})) + O(1/n^2)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \tilde{\lambda}_i(\rho(\mathsf{a}^{(j-1)}) + \Delta\rho(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}))^{\star(i-1)} + O(1/n^2)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \Big(\tilde{\lambda}_i\rho(\mathsf{a}^{(j-1)})^{\star(i-1)}\\
&\quad + \tilde{\lambda}_i(i-1)(\Delta\rho(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})) \star \rho(\mathsf{a}^{(j-1)})^{\star(i-2)}\Big) + O(1/n^2)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \tilde{\lambda}_i\rho(\mathsf{a}^{(j-1)})^{\star(i-1)} + \mathsf{a}_{\mathsf{BMS}} \star (\Delta\rho(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}))\\
&\quad \star \sum_i (i-1)\tilde{\lambda}_i\rho(\mathsf{a}^{(j-1)})^{\star(i-2)} + O(1/n^2)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \tilde{\lambda}(\rho(\mathsf{a}^{(j-1)})) + \mathsf{a}_{\mathsf{BMS}} \star (\Delta\rho(\mathsf{a}^{(j-1)}) - \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}))\\
&\quad \star \lambda'(\rho(\mathsf{a}^{(j-1)})) + O(1/n^2)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \lambda(\rho(\mathsf{a}^{(j-1)})) + \mathsf{a}_{\mathsf{BMS}} \star \Delta\lambda(\rho(\mathsf{a}^{(j-1)})) + \mathsf{a}_{\mathsf{BMS}} \star \Delta\rho(\mathsf{a}^{(j-1)}) \star \lambda'(\rho(\mathsf{a}^{(j-1)}))\\
&\quad - \mathsf{a}_{\mathsf{BMS}} \star \tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)}) \star \lambda'(\rho(\mathsf{a}^{(j-1)})) + O(1/n^2)\\
=&\mathsf{a}^{(j)} + \mathsf{a}_{\mathsf{BMS}} \star \Delta\lambda(\mathsf{b}^{(j)}) + \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \Delta\rho(\mathsf{a}^{(j-1)})\\
&\quad - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})\big) + O(1/n^2).
\end{aligned} \tag{3.40}$$

Note that $\tilde{\Delta}\mathsf{a}^{(0)} = 0$. Since this is of order $O(1/n)$, the recursion confirms that $\tilde{\Delta}\mathsf{a}^{(j)}$ and $\tilde{\Delta}\mathsf{b}^{(j)}$ are of order $O(1/n)$. If we now combine (3.37) and (3.38) with

(3.40), we get

$$\begin{aligned}
\tilde{\Delta}\mathsf{a}^{(j)} &\overset{3.40}{=} -\mathsf{a}_{\mathsf{BMS}} \star \Delta\lambda(\mathsf{b}^{(j)}) - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \Delta\rho(\mathsf{a}^{(j-1)}) \\
&\quad + \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})\big) + O(1/n^2) \\
&\overset{(3.37,3.38)}{=} -\mathsf{a}_{\mathsf{BMS}} \star \left(\frac{(V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\lambda(\mathsf{b}^{(j)}) - (V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{b}^{(j)})}{n\Lambda'(1)}\right) \\
&\quad - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \left(\frac{(C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\rho(\mathsf{a}^{(j-1)}) - (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(j-1)})}{n\Lambda'(1)}\right) \\
&\quad + \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})\big) + O(1/n^2) \\
&= \frac{-1}{n\Lambda'(1)}\Big((V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\mathsf{a}_{\mathsf{BMS}} \star \lambda(\mathsf{b}^{(j)}) - \mathsf{a}_{\mathsf{BMS}} \star (V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{b}^{(j)})\Big) \\
&\quad - \frac{1}{n\Lambda'(1)}\Big((C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \rho(\mathsf{a}^{(j-1)}) \\
&\quad - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(j-1)})\Big) \\
&\quad + \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})\big) + O(1/n^2) \\
&= \frac{-1}{n\Lambda'(1)}\Big((V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\mathsf{a}^{(j)} - \mathsf{a}_{\mathsf{BMS}} \star (V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{b}^{(j)}) \\
&+ (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(1)\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \mathsf{b}^{(j)} - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(j-1)})\Big) \\
&+ \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\tilde{\Delta}\mathsf{a}^{(j-1)} \boxast \rho'(\mathsf{a}^{(j-1)})\big) + O(1/n^2). \qquad (3.41)
\end{aligned}$$

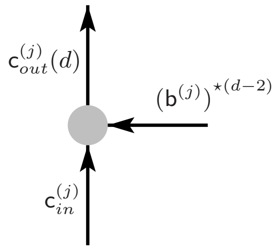
$\mathsf{c}_{out}^{(j)}(d)$
$(\mathsf{b}^{(j)})^{\star(d-2)}$
$\mathsf{c}_{in}^{(j)}$

Figure 3.17: Variable node of degree d.

The preceding formula characterizes the deviation of the densities in $(\mathrm{G}_{\mathrm{T}}^{(\ell)})^c$ due to the realization of $\mathrm{G}_{\mathrm{T}}^{(\ell)}$. As with all previous quantities, let us now write this quantity in a matrix notation. To that end, we need to write in a matrix form the quantity $\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \mathsf{v}$ which appears several times in (3.41) where

v is a vector. Let $\mathsf{c}_{\text{in}}^{(j)}$ be a density. We can write

$$\begin{aligned}\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \mathsf{c}_{\text{in}}^{(j)} &= \sum_d (d-1)\lambda_d \mathsf{a}_{\mathsf{BMS}} \star (\mathsf{b}^{(j)})^{\star(d-2)} \star \mathsf{c}_{\text{in}}^{(j)} \\ &=: \sum_d (d-1)\lambda_d \mathsf{c}_{\text{out}}^{(j)}(d).\end{aligned}$$

where we defined $\mathsf{c}_{\text{out}}^{(j)}(d)$ as the outgoing density at a variable node of degree d which has one incoming density $\mathsf{c}_{\text{in}}^{(j)}$ and $d-2$ incoming densities $\mathsf{b}^{(j)}$ (see Figure 3.17). Let us look at the i^{th} component of this vector. It is equal to

$$\begin{aligned}&\sum_d (d-1)\lambda_d \mathsf{c}_{\text{out}}^{(j)}(d)_i \\ =&\sum_d (d-1)\lambda_d \mathbb{P}\{\nu_{\text{out}}^{(j)} = i \mid \mathrm{d}_v = d\} \\ =&\sum_k \sum_d (d-1)\lambda_d \mathbb{P}\{\nu_{\text{out}}^{(j)} = i \mid \hat{\nu}_{\text{in}}^{(j)} = k, \mathrm{d}_v = d\} \mathbb{P}\{\hat{\nu}_{\text{in}}^{(j)} = k\} \\ =&(\mathrm{V}^{(j)} \mathsf{c}_{\text{in}}^{(j)})_i,\end{aligned}$$

according to the definition of $\mathrm{V}^{(j)}$ in Appendix 3.A.1. In other words, $\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \mathsf{c}_{\text{in}}^{(j)} = \mathrm{V}^{(j)}\mathsf{c}_{\text{in}}^{(j)}$. Similarly, $\rho'(\mathsf{a}^{(j-1)}) \boxast \mathsf{c}_{\text{in}}^{(j-1)} = \mathrm{C}^{(j-1)}\mathsf{c}_{\text{in}}^{(j-1)}$.
Now let us look more precisely at the quantities which appears in (3.41). Fisrt we have $\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \mathsf{b}^{(j)} = \mathrm{V}^{(j)}\mathsf{b}^{(j)}$. Then

$$\begin{aligned}\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(j-1)}) &= \sum_d d C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star (\mathsf{a}^{(j-1)})^{\boxast(d-1)} \\ &= \sum_d d C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} \mathrm{V}^{(j)} (\mathsf{a}^{(j-1)})^{\boxast(d-1)} \\ &= \mathrm{V}^{(j)} (C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(j-1)}).\end{aligned}$$

Finally consider the density $\tilde{\mathsf{a}}^{(j-1)} = \mathsf{a}^{(j-1)} - \tilde{\Delta}\mathsf{a}^{(j-1)}$. We have

$$\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\rho'(\mathsf{a}^{(j-1)}) \boxast \tilde{\mathsf{a}}^{(j-1)}\big) = \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \big(\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)}\big). \tag{3.42}$$

The i^{th} component of the vector $\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)}$ is equal to

$$\begin{aligned}&\big(\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)}\big)_i \\ =&\sum_k \sum_d \rho_d (d-1) \mathbb{P}\{\hat{\nu}_{\text{out}}^{(j)} = i \mid \nu_{\text{in}}^{(j-1)} = k, \mathrm{d}_c = d\} \mathbb{P}\{\nu_{\text{in}}^{(j-1)} = k\} \\ =&\sum_d \rho_d (d-1) \mathbb{P}\{\hat{\nu}_{\text{out}}^{(j)} = i \mid \mathrm{d}_c = d\}.\end{aligned}$$

In other words $\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)} = \sum_d \rho_d(d-1)(\tilde{\mathsf{a}}^{(j-1)})^{\boxast(d-1)}$. So we can write

$$\begin{aligned}
\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \left(\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)}\right) &= \sum_d \rho_d(d-1)\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star (\tilde{\mathsf{a}}^{(j-1)})^{\boxast(d-1)} \\
&= \sum_d \rho_d(d-1)\mathrm{V}^{(j)}(\tilde{\mathsf{a}}^{(j-1)})^{\boxast(d-1)} \\
&= \mathrm{V}^{(j)} \sum_d \rho_d(d-1)(\tilde{\mathsf{a}}^{(j-1)})^{\boxast(d-1)} \\
&= \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)}. \qquad (3.43)
\end{aligned}$$

Combining (3.42) and (3.43), we obtain

$$\begin{aligned}
\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \left(\rho'(\mathsf{a}^{(j-1)}) \boxast \tilde{\mathsf{a}}^{(j-1)}\right) =& \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\tilde{\mathsf{a}}^{(j-1)} \qquad (3.44) \\
=& \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathsf{a}^{(j-1)} + \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\tilde{\Delta}\mathsf{a}^{(j-1)}.
\end{aligned}$$

According to (3.44), if we replace $\tilde{\mathsf{a}}^{(j-1)}$ by $\mathsf{a}^{(j-1)}$, we have that $\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \left(\rho'(\mathsf{a}^{(j-1)}) \boxast \mathsf{a}^{(j-1)}\right) = \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathsf{a}^{(j-1)}$. We can then conclude that

$$\mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(j)}) \star \left(\rho'(\mathsf{a}^{(j-1)}) \boxast \tilde{\Delta}\mathsf{a}^{(j-1)}\right) = \mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\tilde{\Delta}\mathsf{a}^{(j-1)}.$$

Moreover, we have that

$$\begin{aligned}
\mathsf{a}_{\mathsf{BMS}} \star (V^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{b}^{(\ell)}) &= \sum_d dV_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} \mathsf{a}_{\mathsf{BMS}} \star (\mathsf{b}^{(\ell)})^{\star(d-1)} \\
&= \sum_d dV_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} \mathsf{a}^{(\ell)}(d),
\end{aligned}$$

where $\mathsf{a}^{(\ell)}(d)$ is the outgoing density at a variable node of degree d in iteration ℓ. In the same way we can write

$$(C^{\mathrm{G}_{\mathrm{T}}^{(\ell)}})'(\mathsf{a}^{(\ell-1)}) = \sum_d dC_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} \mathsf{b}^{(\ell)}(d).$$

We can thus write $\tilde{\Delta}\mathsf{a}^{(\ell)}$ as

$$\begin{aligned}\tilde{\Delta}\mathsf{a}^{(\ell)} =& \frac{-1}{n\Lambda'(1)}\big((V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\mathsf{a}^{(\ell)} - \sum_{d} dV_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\mathsf{a}^{(\ell)}(d)\\ &+ (C^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\mathrm{V}^{(\ell)}\mathsf{b}^{(\ell)} - \mathrm{V}^{(\ell)}\sum_{d} dC_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\mathsf{b}^{(\ell)}(d)\big)\\ &+ \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell-1)}\tilde{\Delta}\mathsf{a}^{(\ell-1)} + O(1/n^2)\\ =& \frac{-1}{n\Lambda'(1)}\Big(\sum_{d} dV_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{a}^{(\ell)} - \mathsf{a}^{(\ell)}(d)) + \mathrm{V}^{(\ell)}\sum_{d} dC_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{b}^{(\ell)} - \mathsf{b}^{(\ell)}(d))\Big)\\ &+ \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell-1)}\tilde{\Delta}\mathsf{a}^{(\ell-1)} + O(1/n^2)\\ =& \frac{-1}{n\Lambda'(1)}\sum_{i=1}^{\ell}\prod_{k=\ell}^{i+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\big(\sum_{d} dV_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{a}^{(i)} - \mathsf{a}^{(i)}(d))\\ &+ \mathrm{V}^{(i)}\sum_{d} dC_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{b}^{(i)} - \mathsf{b}^{(i)}(d))\big) + O(1/n^2). \qquad (3.45)\end{aligned}$$

Messages Correction

The second term which gives a non-vanishing contribution in the irregular case is due to messages that flow across the boundary from $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ to $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$. These messages influence the neighbors of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ and thus the average densities in $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$. Indeed, assume that the density of messages which flow from $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ across the boundary to $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ at time $j < \ell$ is not $\mathsf{a}^{(j)}$ but $\mathsf{a}_{\star}^{(j)}$. This influences messages up to a distance $\ell - j$ away from the boundary.

More precisely, consider an outgoing edge at a variable node at a distance j from the boundary of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$. Let $\hat{\mathsf{a}}_j^{(\ell)}$ be its message density at iteration ℓ and let n_j be the number of such edges. In the same way define $\hat{\mathsf{b}}_{\ell}^{(j)}$ and m_j for an outgoing edge at a check node. Thus, if we pick uniformly at random an edge from $(\mathrm{G}_{\mathtt{T}}^{(\ell)})^c$ its message density is

$$\begin{aligned}\hat{\mathsf{a}}^{(\ell)} &= \frac{(n\Lambda'(1) - |\mathrm{G}_{\mathtt{T}}^{(\ell)}| - \sum_{j=1}^{\ell-1} n_j)\mathsf{a}^{(\ell)} + \sum_{j=1}^{\ell-1} n_j(\mathsf{a}_j^{(\ell)} - \hat{\Delta}\mathsf{a}_j^{(\ell)})}{n\Lambda'(1) - |\mathrm{G}_{\mathtt{T}}^{(\ell)}|}\\ &= \mathsf{a}^{(\ell)} - \underbrace{\frac{1}{n\Lambda'(1) - |\mathrm{G}_{\mathtt{T}}^{(\ell)}|}\sum_{j=1}^{\ell-1} n_j\hat{\Delta}\mathsf{a}_j^{(\ell)}}_{\hat{\Delta}\mathsf{a}^{(\ell)}}. \qquad (3.46)\end{aligned}$$

In order to compute $n_j\hat{\Delta}\mathsf{a}_j^{(\ell)}$, we can expand $n_j\hat{\mathsf{a}}_j^{(\ell)}$ as follows

$$
\begin{aligned}
&n_j\hat{\mathsf{a}}_j^{(\ell)}\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \lambda_i (i-1)(\mathsf{b}^{(\ell)})^{\star(i-2)} \star m_j \hat{\mathsf{b}}_j^{(\ell)}\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \lambda_i (i-1)(\mathsf{b}^{(\ell)})^{\star(i-2)} \star \Big(\sum_i \rho_i(i-1)(\mathsf{a}^{(\ell-1)})^{\boxast(i-2)} \boxast n_{j-1}\hat{\mathsf{a}}_{j-1}^{(\ell-1)}\Big)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \lambda_i (i-1)(\mathsf{b}^{(\ell)})^{\star(i-2)} \star \Big(\sum_i \rho_i(i-1)(\mathsf{a}^{(\ell-1)})^{\boxast(i-2)} \boxast n_{j-1}\big(\mathsf{a}^{(\ell-1)} - \hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\big)\Big)\\
=&\mathsf{a}_{\mathsf{BMS}} \star \sum_i \lambda_i (i-1)(\mathsf{b}^{(\ell)})^{\star(i-2)} \star \Big(m_j \mathsf{b}^{(\ell)} - \sum_i \rho_i(i-1)(\mathsf{a}^{(\ell-1)})^{\boxast(i-2)} \boxast n_{j-1}\hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\Big)\\
=&n_j\mathsf{a}^{(\ell)} - \mathsf{a}_{\mathsf{BMS}} \star \sum_i \lambda_i (i-1)(\mathsf{b}^{(\ell)})^{\star(i-2)} \star \Big(\sum_i \rho_i(i-1)(\mathsf{a}^{(\ell-1)})^{\boxast(i-2)} \boxast n_{j-1}\hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\Big)\\
=&n_j\mathsf{a}^{(\ell)} - \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(\ell)}) \star \big(\rho'(\mathsf{a}^{(\ell-1)}) \boxast n_{j-1}\hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\big).
\end{aligned}
$$

Thus we can derive $\hat{\Delta}\mathsf{a}_j^{(\ell)}$ from the recursion

$$
n_j\hat{\Delta}\mathsf{a}_j^{(\ell)} = \mathsf{a}_{\mathsf{BMS}} \star \lambda'(\mathsf{b}^{(\ell)}) \star \big(\rho'(\mathsf{a}^{(\ell-1)}) \boxast n_{j-1}\hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\big),
$$

with

$$
n_0\hat{\Delta}\mathsf{a}_0^{(\ell-j)} = (|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)} - \mathsf{a}_\star^{(\ell-j)}).
$$

As we did for the degree distribution correction, we can rewrite $n_j\hat{\Delta}\mathsf{a}_j^{(\ell)}$ in a matrix form as

$$
\begin{aligned}
n_j\hat{\Delta}\mathsf{a}_j^{(\ell)} &= \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell-1)} n_{j-1}\hat{\Delta}\mathsf{a}_{j-1}^{(\ell-1)}\\
&= \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell-1)} \cdots \mathrm{V}^{(\ell-j+1)}\mathrm{C}^{(\ell-j)} n_0\hat{\Delta}\mathsf{a}_0^{(\ell-j)}\\
&= \prod_{k=\ell}^{\ell-j+1} \mathrm{V}^{(k)} C^{(k-1)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)} - \mathsf{a}_\star^{(\ell-j)}).
\end{aligned}
$$

Thus the deviation due to the messages correction can be written as

$$
\hat{\Delta}\mathsf{a}^{(\ell)} = \frac{1}{n\Lambda'(1)- \mid \mathrm{G}_{\mathtt{T}}^{(\ell)} \mid} \sum_{j=1}^{\ell-1} \prod_{k=\ell}^{\ell-j+1} \mathrm{V}^{(k)} C^{(k-1)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)} - \mathsf{a}_\star^{(\ell-j)}). \tag{3.47}
$$

Putting it Together

In order to derive S^c we need to compute for $i \in (\mathtt{T}^{(\ell)})^c$ the expectation of $\mu_i^{(\ell)}$ for a particular realization of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$, i.e., $\mathbb{E}[\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]$ (see (3.36)). According to the two corrections, we know that the variable-to-check density on an edge picked uniformly at random from $(\mathtt{T}^{(\ell)})^c$ is equal to $\mathsf{a}^{(\ell)} - \tilde{\Delta}\mathsf{a}^{(\ell)} - \hat{\Delta}\mathsf{a}^{(\ell)}$. Thus we can write

$$\begin{aligned}
\sum_{i\in(\mathtt{T}^{(\ell)})^c} \mathbb{E}[\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}] &= |(\mathtt{T}^{(\ell)})^c|\mathfrak{E}\big(\mathsf{a}^{(\ell)} - \tilde{\Delta}\mathsf{a}^{(\ell)} - \hat{\Delta}\mathsf{a}^{(\ell)}\big) \\
&= |(\mathtt{T}^{(\ell)})^c|\Big(\mathfrak{E}\big(\mathsf{a}^{(\ell)}\big) - \mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big) - \mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\Big) \\
&= |(\mathtt{T}^{(\ell)})^c|\Big(x^{(\ell)} - \mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big) - \mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\Big). \qquad (3.48)
\end{aligned}$$

We can now expand (3.36) as

$$\begin{aligned}
S^c &= \lim_{n\to\infty}\Big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\Big[\mathbb{E}\big[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\big] \sum_{i\in(\mathtt{T}^{(\ell)})^c} \mathbb{E}\big[\mu_i^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\big]\Big] - n\Lambda'(1)(x^{(\ell)})^2\Big) \\
&\overset{(3.48)}{=} \lim_{n\to\infty}\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[|(\mathtt{T}^{(\ell)})^c|\mathbb{E}\big[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\big]\big(x^{(\ell)} - \mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big)\text{-}\mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\big)\big]\text{-}n\Lambda'(1)(x^{(\ell)})^2\big) \\
&= \lim_{n\to\infty}\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[\big(n\Lambda'(1)\text{-}|\mathtt{T}^{(\ell)}|\big)\mathbb{E}\big[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\big]\big(x^{(\ell)}\text{-}\mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big)\text{-}\mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\big)\big]\text{-}n\Lambda'(1)(x^{(\ell)})^2\big) \\
&= \lim_{n\to\infty} n\Lambda'(1)\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]x^{(\ell)}\big] - (x^{(\ell)})^2\big) \\
&\quad - \lim_{n\to\infty} n\Lambda'(1)\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[\mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big)\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]\big] + \mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[\mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]\big]\big) \\
&\quad - \lim_{n\to\infty} \mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[|\mathtt{T}^{(\ell)}|\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]x^{(\ell)}\big] \\
&\quad + \lim_{n\to\infty}\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[|\mathtt{T}^{(\ell)}|\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}](\mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big) + \mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big))\big]\big) \\
&\overset{\text{(i)}}{=} \lim_{n\to\infty} n\Lambda'(1)\underbrace{\big(x^{(\ell)}\overbrace{\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]\big]}^{x^{(\ell)}} - (x^{(\ell)})^2\big)}_{=0} \\
&\quad - \lim_{n\to\infty} n\Lambda'(1)\big(\mathbb{E}\big[\mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big)\mu_1^{(\ell)}\big] + \mathbb{E}\big[\mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big)\mu_1^{(\ell)}\big]\big) \\
&\quad - \lim_{n\to\infty} x^{(\ell)}\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[|\mathtt{T}^{(\ell)}|\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}]\big] \\
&\overset{\text{(ii)}}{=} - \lim_{n\to\infty} n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}\big[\tilde{\Delta}\mathsf{a}^{(\ell)}\mu_1^{(\ell)}\big]\big) \qquad (3.49) \\
&\quad - \lim_{n\to\infty} n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}\big[\hat{\Delta}\mathsf{a}^{(\ell)}\mu_1^{(\ell)}\big]\big) \qquad (3.50) \\
&\quad - \lim_{n\to\infty} x^{(\ell)}\mathbb{E}\big[(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\mu_1^{(\ell)}\big],
\end{aligned}$$

where in step (i) $\lim_{n\to\infty}\big(\mathbb{E}_{\mathrm{G}_{\mathtt{T}}^{(\ell)}}\big[|\mathtt{T}^{(\ell)}|\mathbb{E}[\mu_1^{(\ell)} \mid \mathrm{G}_{\mathtt{T}}^{(\ell)}](\mathfrak{E}\big(\tilde{\Delta}\mathsf{a}^{(\ell)}\big) + \mathfrak{E}\big(\hat{\Delta}\mathsf{a}^{(\ell)}\big))\big]\big)=0$, since $\tilde{\Delta}\mathsf{a}^{(\ell)}$ and $\hat{\Delta}\mathsf{a}^{(\ell)}$ are of order $1/n$, and in step (ii) we replaced $|\mathtt{T}^{(\ell)}|$ by $(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)$. According to the derivations in Sections 3.A.5 and 3.A.5, we are now able to expand (3.49) as

$$\begin{aligned}
&\lim_{n\to\infty} n\Lambda'(1)\mathbb{E}\big[\mu_1^{(\ell)}\tilde{\Delta}\mathsf{a}^{(\ell)}\big]\\
&= \lim_{n\to\infty} n\Lambda'(1)\mathbb{E}\Big[\frac{-\mu_1^{(\ell)}}{n\Lambda'(1)}\sum_{i=1}^{\ell}\prod_{k=\ell}^{i+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\Big(\sum_d dV_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{a}^{(i)}-\mathsf{a}^{(i)}(d))\\
&\qquad\qquad + \mathrm{V}^{(i)}\sum_d dC_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{b}^{(i)}-\mathsf{b}^{(i)}(d))\Big)\Big] + O(1/n^2)\\
&=\mathbb{E}\Big[-\mu_1^{(\ell)}\sum_{i=1}^{\ell}\prod_{k=\ell}^{i+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\Big(\sum_d dV_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{a}^{(i)}-\mathsf{a}^{(i)}(d))\\
&\qquad\qquad + \mathrm{V}^{(i)}\sum_d dC_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(\mathsf{b}^{(i)}-\mathsf{b}^{(i)}(d))\Big)\Big]\\
&= -\sum_{i=1}^{\ell}\prod_{k=\ell}^{i+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\Big(\sum_d d\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}](\mathsf{a}^{(i)}-\mathsf{a}^{(i)}(d))\\
&\qquad\qquad + \mathrm{V}^{(i)}\sum_d d\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}](\mathsf{b}^{(i)}-\mathsf{b}^{(i)}(d))\Big),
\end{aligned}$$

and (3.50) as

$$\begin{aligned}
&\lim_{n\to\infty} n\Lambda'(1)\mathbb{E}\big[\mu_1^{(\ell)}\hat{\Delta}\mathsf{a}^{(\ell)}\big]\\
&= \lim_{n\to\infty} n\Lambda'(1)\mathbb{E}\Big[\frac{\mu_1^{(\ell)}}{n\Lambda'(1)-\mid \mathrm{G}_{\mathtt{T}}^{(\ell)}\mid}\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}V^{(k)}C(k-1)(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)}-\mathsf{a}_\star^{(\ell-j)})\Big]\\
&=\mathbb{E}\Big[\mu_1^{(\ell)}\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}\mathrm{V}^{(k)}C(k-1)(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)}-\mathsf{a}_\star^{(\ell-j)})\Big]\\
&=\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)}-\mathsf{a}_\star^{(\ell-j)})\big]\\
&=\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)(\mathsf{a}^{(\ell-j)}-\mathsf{a}_\star^{(\ell-j)})\big]\\
&=\sum_{j=1}^{\ell-1}\prod_{k=\ell}^{\ell-j+1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\Big(\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\big]\mathsf{a}^{(\ell-j)}-\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}\big]\Big)
\end{aligned}$$

Finally it remains to expand $\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}]$, $\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}]$, $\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\big]$ and $\mathbb{E}\big[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}\big]$ in order to be able to compute S^c. Let us look at each of these terms separately.

$\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}]$

Let us define $V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j)$ as the number of variable nodes of degree d at distance j of the root in the future of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ if j is positive and in the past of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ if j is negative. Thus we have

$$V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}} = \sum_{j=1}^{\ell} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j) + \sum_{j=0}^{2\ell} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)$$

Then we can write

$$\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}] &= \mathbb{E}\Big[\mu_1^{(\ell)}\Big(\sum_{j=1}^{\ell} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j) + \sum_{j=0}^{2\ell} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)\Big)\Big] \\
&= \sum_{j=1}^{\ell} \mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j)] + \sum_{j=0}^{2\ell} \mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)].
\end{aligned}$$

First we compute

$$\begin{aligned}
\sum_{j=1}^{\ell} \mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j)] &= \mathbb{E}[\mu_1^{(\ell)}] \sum_{j=1}^{\ell} \mathbb{E}[V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(j)] \\
&= x^{(\ell)} \sum_{j=1}^{\ell} \rho'(1)^j \lambda'(1)^{j-1} \lambda_d.
\end{aligned}$$

Secondly, we can expand $\sum_{j=0}^{2\ell} \mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)]$ in the same way we expand $\mathbb{E}[|\mathsf{B}_{3,k}|\mu_1^{(\ell)} \mid \nu_{i_1}^{(\ell-k)} = r]$ in Appendix 3.A.3. If $j \le \ell$

$$\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)] = \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \lambda_d \mathsf{a}^{(\ell-j)}(d).$$

Further, note that if $j > \ell$, then $\mathbb{E}[V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)] = \mathbb{E}[|\mathsf{B}_{3,\ell}|(\lambda'(1)\rho'(1))^{j-\ell}\lambda_d]$. We can thus write if $j > \ell$

$$\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-j)] &= \mathbb{E}[\mu_1^{(\ell)} |\mathsf{B}_{3,\ell}|(\lambda'(1)\rho'(1))^{j-\ell}\lambda_d] \\
&= \mathbb{E}[\mu_1^{(\ell)} |\mathsf{B}_{3,\ell}|](\lambda'(1)\rho'(1))^{j-\ell}\lambda_d \\
&= \prod_{k=\ell}^{1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathsf{a}_{\mathsf{BMS}} (\lambda'(1)\rho'(1))^{j-\ell}\lambda_d.
\end{aligned}$$

Combining everything, we obtain

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}] =& x^{(\ell)} \sum_{j=1}^{\ell} \rho'(1)^j \lambda'(1)^{j-1} \lambda_d \\ &+ \sum_{j=0}^{\ell} \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \lambda_d \mathsf{a}^{(\ell-j)}(d) \\ &+ \sum_{j=\ell+1}^{2\ell} \mathbf{e}_m^T \prod_{k=\ell}^{1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathsf{a}_{\mathsf{BMS}} (\lambda'(1)\rho'(1))^{j-\ell} \lambda_d. \qquad (3.51)\end{aligned}$$

$\mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}]$

The computation is similar to the previous one. First let us define $C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j)$ as the number of check nodes of degree d at distance j of the root in the future of $\mathrm{G}_{\mathrm{T}}^{(\ell)}$ if j is positive and in the past of $\mathrm{G}_{\mathrm{T}}^{(\ell)}$ if j is negative. Thus we have

$$C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}} = \sum_{j=0}^{\ell-1} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j) + \sum_{j=1}^{2\ell} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j).$$

Similarly to what we did above, we expand

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}] &= \mathbb{E}\Big[\mu_1^{(\ell)} \Big(\sum_{j=0}^{\ell-1} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j) + \sum_{j=1}^{2\ell} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j) \Big)\Big] \\ &= \sum_{j=0}^{\ell-1} \mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j)] + \sum_{j=1}^{2\ell} \mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j)].\end{aligned}$$

First we compute

$$\begin{aligned}\sum_{j=0}^{\ell-1} \mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j)] &= \mathbb{E}[\mu_1^{(\ell)}] \sum_{j=0}^{\ell-1} \mathbb{E}[C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(j)] \\ &= x^{(\ell)} \sum_{j=0}^{\ell-1} (\rho'(1)\lambda'(1))^j \rho_d.\end{aligned}$$

Secondly, we look at the second term $\sum_{j=1}^{2\ell} \mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j)]$. Again we have to distinguish between two cases. If $j \leq \ell$

$$\mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j)] = \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+2} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathrm{V}^{(\ell-j+1)} \rho_d \mathsf{b}^{(\ell-j+1)}(d),$$

and if $j > \ell$

$$\mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-j)] = \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+2} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathrm{V}^{(\ell-j+1)} \mathbf{e}_0 (\lambda'(1)\rho'(1))^{j-\ell-1} \rho_d.$$

Combining everything, we obtain

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)} C_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}] =& x^{(\ell)} \sum_{j=0}^{\ell-1} (\rho'(1)\lambda'(1))^j \rho_d \\ &+ \sum_{j=1}^{\ell} \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+2} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathrm{V}^{(\ell-j+1)} \rho_d \mathsf{b}^{(\ell-j+1)}(d) \\ &+ \sum_{j=\ell+1}^{2\ell} \mathbf{e}_m^T \prod_{k=\ell}^{\ell-j+2} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathrm{V}^{(\ell-j+1)} \mathbf{e}_0 (\lambda'(1)\rho'(1))^{j-\ell-1} \rho_d.\end{aligned} \tag{3.52}$$

$\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)]$

First note that

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)] &= \mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{1,\ell}|] + \mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|] \\ &= \underbrace{\mathbb{E}[\mu_1^{(\ell)}]}_{x^{(\ell)}} \underbrace{\mathbb{E}[|\mathsf{B}_{1,\ell}|]}_{(\lambda'(1)\rho'(1))^{\ell}} + \mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|].\end{aligned}$$

Let us rewrite $\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|]$ as

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|] &= \mathbb{E}\Big[\mu_1^{(\ell)} \sum_d (d-1) V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-2\ell)\Big] \\ &= \sum_d (d-1) \mathbb{E}\Big[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-2\ell)\Big].\end{aligned}$$

The computation of $\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-2\ell)]$ has already been performed in Appendix 3.A.5. So we have

$$\begin{aligned}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|] &= \sum_d (d-1)\lambda_d (\lambda'(1)\rho'(1))^{\ell} \mathbf{e}_m^T \prod_{k=\ell}^{1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathsf{a}_{\mathsf{BMS}} \\ &= \lambda'(1)(\lambda'(1)\rho'(1))^{\ell} \mathbf{e}_m^T \prod_{k=\ell}^{1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathsf{a}_{\mathsf{BMS}}.\end{aligned}$$

We can thus write

$$\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)] = (\lambda'(1)\rho'(1))^{\ell}\Big(x^{(\ell)} + \lambda'(1)\mathbf{e}_m^T \prod_{k=\ell}^{1} \mathrm{V}^{(k)} \mathrm{C}^{(k-1)} \mathsf{a}_{\mathsf{BMS}}\Big).$$

$\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}]$

$$\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}|+|\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}] &= \mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{1,\ell}|\mathsf{a}_\star^{(\ell-j)}]+\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)}]\\
&= \mathbb{E}[\mu_1^{(\ell)}]\mathbb{E}[|\mathsf{B}_{1,\ell}|\mathsf{a}_\star^{(\ell-j)}]+\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)}]\\
&= x^{(\ell)}\mathbb{E}[|\mathsf{B}_{1,\ell}|\mathsf{a}_\star^{(\ell-j)}]+\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)}].
\end{aligned}$$

Now let us make the following observation. First let us number the edges in $\mathsf{B}_{4,\ell}$ from i_1 to $i_{|\mathsf{B}_{4,\ell}|}$. Then for all edges $i_j \in \mathsf{B}_{4,\ell}$, $j \in \{1,\dots,|\mathsf{B}_{4,\ell}|\}$, let us call B_{i_j} the boundary of its support tree of edge height ℓ. We can thus write $|\mathsf{B}_{4,2\ell}|=\sum_{j=1}^{|\mathsf{B}_{4,\ell}|}|B_{i_j}|$. So we can write

$$\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)}] &= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\big[\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|]\big]\\
&= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\big[\mathbb{E}[\mu_1^{(\ell)} \mid |\mathsf{B}_{4,\ell}|]\mathbb{E}[|\mathsf{B}_{4,2\ell}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|]\big]\\
&= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\Big[\mathbb{E}[\mu_1^{(\ell)} \mid |\mathsf{B}_{4,\ell}|]\mathbb{E}\Big[\sum_{j=1}^{|\mathsf{B}_{4,\ell}|}|B_{i_j}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|\Big]\Big]\\
&= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\Big[\mathbb{E}[\mu_1^{(\ell)} \mid |\mathsf{B}_{4,\ell}|]\sum_{j=1}^{|\mathsf{B}_{4,\ell}|}\mathbb{E}[|B_{i_j}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|]\Big]\\
&= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\big[\mathbb{E}[\mu_1^{(\ell)} \mid |\mathsf{B}_{4,\ell}|]|\mathsf{B}_{4,\ell}|\mathbb{E}[|B_{i_1}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|]\big]\\
&= \mathbb{E}_{|\mathsf{B}_{4,\ell}|}\big[\mathbb{E}[\mu_1^{(\ell)} \mid |\mathsf{B}_{4,\ell}|]|\mathsf{B}_{4,\ell}|\big]\mathbb{E}_{|\mathsf{B}_{4,\ell}|}\big[\mathbb{E}[|B_{i_1}|\mathsf{a}_\star^{(\ell-j)} \mid |\mathsf{B}_{4,\ell}|]\big]\\
&= \mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,\ell}|]\mathbb{E}[|B_{i_1}|\mathsf{a}_\star^{(\ell-j)}].
\end{aligned}$$

First we compute

$$\begin{aligned}
\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,\ell}|] &= \mathbb{E}\big[\mu_1^{(\ell)}\sum_d (d-1)V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-\ell)\big]\\
&= \sum_d (d-1)\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\ell)}}(-\ell)]\\
&\overset{\text{(i)}}{=} \sum_d (d-1)\lambda_d \mathbf{e}_m^T \prod_{k=\ell}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}\\
&= \lambda'(1)\mathbf{e}_m^T \prod_{k=\ell}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}},
\end{aligned}$$

where the derivation of $\mathbb{E}[\mu_1^{(\ell)} V_d^{\mathrm{G}_{\mathtt{T}}^{(\ell)}}(-\ell)]$ in step (i) has been performed in Appendix 3.A.5. We still need to compute

$$\begin{aligned}\mathbb{E}[|\mathsf{B}_{1,\ell}|\mathsf{a}_\star^{(\ell-j)}] &= \mathrm{V}^{(\ell)}\mathrm{C}^{(\ell)}\mathbb{E}[|\mathsf{B}_{1,\ell-1}|\mathsf{a}_\star^{(\ell-j-1)}]\\ &= \prod_{k=\ell-j}^{1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathbb{E}[|\mathsf{B}_{1,j}|\mathsf{a}_\star^{(0)}]\\ &= \prod_{k=\ell-j}^{1} \mathrm{V}^{(k)}\mathrm{C}^{(k-1)}(\lambda'(1)\rho'(1))^j\mathsf{a}_{\mathsf{BMS}}.\end{aligned}$$

Finally we note that $\mathbb{E}[|B_{i_1}|\mathsf{a}_\star^{(\ell-j)}] = \mathbb{E}[|\mathsf{B}_{1,\ell}|\mathsf{a}_\star^{(\ell-j)}]$ for all $i_j \in \mathtt{B}_{4,\ell}$. We can thus write

$$\begin{aligned}&\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\ell}| + |\mathsf{B}_{4,2\ell}|)\mathsf{a}_\star^{(\ell-j)}]\\ =&\Big(x^{(\ell)} + \lambda'(1)\mathbf{e}_m^T\prod_{k=\ell}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}\Big)(\lambda'(1)\rho'(1))^j\prod_{k=\ell-j}^{1}\mathrm{V}^{(k)}\mathrm{C}^{(k-1)}\mathsf{a}_{\mathsf{BMS}}.\end{aligned}$$

3.B Details of Asymptotic Variance Computation

3.B.1 Edges in $\mathtt{T}_1$

$$\begin{aligned}&\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_1^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\ =&\lim_{\ell\to\infty}\lim_{n\to\infty}\sum_{k=1}^{\mathbb{k}}\sum_{r=-m}^{m}\mathbf{u}_m^T\prod_{j=\ell}^{\ell-k+1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{u}_r\ \mathbb{P}\{\nu_1^{(\ell)}=m,\nu_1^{(\ell-k)}=r\}\\ =&\mathbf{u}_m^T\sum_{k=1}^{\mathbb{k}}(\mathrm{VC})^k\mathrm{P}_m\\ =&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}(\mathrm{I}-(\mathrm{VC})^{\mathbb{k}})\mathrm{P}_m.\end{aligned}\tag{3.53}$$

We can then write (3.53) as

$$\begin{aligned}=&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m - \mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}(\mathrm{VC})^{\mathbb{k}}\sum_{i=m}^{2m+1}p_i\mathbf{e}_i\\ =&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m - \mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\sum_{i=m}^{2m+1}\lambda_i^{\mathbb{k}}p_i\mathbf{e}_i\end{aligned}$$

3.B.2 Edges in $\mathtt{T}_2$

$$
\begin{aligned}
\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_2^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]=&\lim_{\ell\to\infty}\lim_{n\to\infty}(x^{(\ell)})^2\sum_{k=1}^{\Bbbk}\rho'(1)^k\lambda'(1)^{k-1}\\
=&x^2\rho'(1)\frac{1-(\lambda'(1)\rho'(1))^{\Bbbk}}{1-\lambda'(1)\rho'(1)}.
\end{aligned}
\tag{3.54}
$$

3.B.3 Edges in $\mathtt{T}_3$

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty}\sum_{k=1}^{\ell}\mathbb{E}\Big[\sum_{i\in\mathtt{B}_{3,k}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
\overset{(i)}{=}&\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbf{u}_m^T\sum_{k=1}^{2\Bbbk}\sum_{r=-m}^{m}\prod_{j=\ell}^{\ell-k+1}\mathrm{V}^{(j)}\mathrm{C}^{(j-1)}\mathbf{u}_r\mathbb{P}\{\nu^{(\ell)}=m,\nu^{(\ell-k)}=r\}\\
=&\mathbf{u}_m^T\sum_{k=1}^{2\Bbbk}(\mathrm{VC})^k\mathrm{P}_m\\
=&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}(\mathrm{I}-(\mathrm{VC})^{2\Bbbk})\mathrm{P}_m\\
=&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\sum_{i=0}^{2m+1}\lambda_i^{2\Bbbk}p_i\mathbf{e}_i.
\end{aligned}
$$

3.B.4 Edges in $\mathtt{T}_4$

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_4^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
&=\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{l=0}^{2\Bbbk}\sum_{i\in\mathtt{B}_{4,l}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
&=\lim_{\ell\to\infty}\lim_{n\to\infty}\left(\sum_{h=0}^{\Bbbk}\Big(\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,h}\Big)^T\mathrm{F}\mathrm{B}^{2h,0}\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h,j}\mathrm{B}^{2h,j+1}\mathbf{c}^{2h,h}\right.\\
&\quad+\sum_{h=0}^{\Bbbk-1}\Big(\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,h}\Big)^T\\
&\qquad\qquad\left.\big(\mathrm{B}^{2h+1,0}\big)^T\mathrm{F}\hat{\mathrm{B}}^{2h+1,\text{-}1}\mathrm{B}^{2h+1,0}\prod_{j=0}^{h\text{-}1}\hat{\mathrm{B}}^{2h+1,j}\mathrm{B}^{2h+1,j+1}\mathbf{c}^{2h+1,h}\right)\\
&=\lim_{n\to\infty}\left(\sum_{h=0}^{\Bbbk}\Big((\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}\Big)^T\mathrm{F}\mathrm{B}_0(\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}+\sum_{h=0}^{\Bbbk-1}\Big((\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}\Big)^T\mathrm{B}^T\mathrm{F}\hat{\mathrm{B}}_0\mathrm{B}(\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}\right)\\
&=\sum_{h=0}^{\Bbbk-1}((\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c})^T\mathrm{K}(\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}+((\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c})^T\mathrm{F}\mathrm{B}_0(\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c}.
\end{aligned}\tag{3.55}
$$

where in the last step we define $\mathrm{K}=\mathrm{F}\mathrm{B}_0+\mathrm{B}^T\mathrm{F}\hat{\mathrm{B}}_0\mathrm{B}$ in order to simplify the notation.

We can then rewrite (3.55) as:

$$
\begin{aligned}
&\sum_{h=0}^{\mathbb{k}-1}((\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c})^T\mathrm{K}(\hat{\mathrm{B}}\mathrm{B})^h\mathbf{c}+((\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\mathbf{c})^T\mathrm{FB}_0(\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\mathbf{c}\\
=&\sum_{h=0}^{\mathbb{k}-1}((\hat{\mathrm{B}}\mathrm{B})^h\sum_{i=0}^{(2m+1)^2}c_i\tilde{\mathbf{e}}_i\mathrm{K}(\hat{\mathrm{B}}\mathrm{B})^h\sum_{i=0}^{(2m+1)^2}c_i\tilde{\mathbf{e}}_i\\
&+((\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\sum_{i=0}^{(2m+1)^2}c_i\tilde{\mathbf{e}}_i\mathrm{FB}_0(\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\sum_{i=0}^{(2m+1)^2}c_i\tilde{\mathbf{e}}_i\\
=&\sum_{h=0}^{\mathbb{k}-1}\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^hc_i\tilde{\mathbf{e}}_i+h\tilde{\lambda}_2^{h-1}c_3\tilde{\mathbf{e}}_2\Big)^T\mathrm{K}\Big(\sum_{i=1}^{2m+1}\tilde{\lambda}_i^hc_i\tilde{\mathbf{e}}_i+h\tilde{\lambda}_2^{h-1}c_3\tilde{\mathbf{e}}_2\Big)\\
&\qquad+\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}c_i\tilde{\mathbf{e}}_i+\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}c_3\tilde{\mathbf{e}}_2\Big)^T\mathrm{FB}_0\Big(\sum_{i=1}^{2m+1}\tilde{\lambda}_i^{\mathbb{k}}c_i\tilde{\mathbf{e}}_i+\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}c_3\tilde{\mathbf{e}}_2\Big)\\
=&\sum_{h=0}^{\mathbb{k}-1}\Big(\sum_{i,j=1}^{(2m+1)^2}\tilde{\lambda}_i^h\tilde{\lambda}_j^hc_ic_j\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_j+h^2\tilde{\lambda}_2^{2(h-1)}c_3^2\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_2\\
&\qquad+\sum_{i=1}^{(2m+1)^2}h\tilde{\lambda}_2^{h-1}\tilde{\lambda}_i^hc_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{K}\mathbf{e}_i+\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_2)\Big)\\
&+\sum_{i,j=1}^{(2m+1)^2}\lambda_i^{\mathbb{k}}\tilde{\lambda}_j^{\mathbb{k}}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{FD}\tilde{\mathbf{e}}_j+\mathbb{k}^2\tilde{\lambda}_2^{2(\mathbb{k}-1)}c_3^2\tilde{\mathbf{e}}_2^T\mathrm{FD}\tilde{\mathbf{e}}_2\\
&+\sum_{i=1}^{(2m+1)^2}\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}\tilde{\lambda}_i^{\mathbb{k}}c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{FD}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{FD}\tilde{\mathbf{e}}_2)\\
=&\sum_{i,j=1}^{(2m+1)^2}\frac{1-(\tilde{\lambda}_i\tilde{\lambda}_j)^{\mathbb{k}}}{1-\tilde{\lambda}_i\tilde{\lambda}_j}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_j\\
&+\frac{1+\tilde{\lambda}_2^2-\tilde{\lambda}_2^{2\mathbb{k}}-\tilde{\lambda}_2^{2(\mathbb{k}-1)}(\mathbb{k}+\tilde{\lambda}_2^2-\mathbb{k}\tilde{\lambda}_2^2)}{(1-\tilde{\lambda}_2^2)^3}c_3^2\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_2\\
&+\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i\Big(\frac{1-(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}}}{(1-\tilde{\lambda}_2\tilde{\lambda}_i)^2}-\frac{\mathbb{k}(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}-1}}{1-\tilde{\lambda}_2\tilde{\lambda}_i}\Big)c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_2)\\
&+\sum_{i,j=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}\tilde{\lambda}_j^{\mathbb{k}}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{FD}\tilde{\mathbf{e}}_j+\mathbb{k}^2\tilde{\lambda}_2^{2(\mathbb{k}-1)}c_3^2\tilde{\mathbf{e}}_2^T\mathrm{FD}\tilde{\mathbf{e}}_2\\
&+\sum_{i=1}^{(2m+1)^2}\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}\tilde{\lambda}_i^{\mathbb{k}}c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{FD}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{FD}\tilde{\mathbf{e}}_2). \qquad (3.56)
\end{aligned}
$$

We can thus finaly write,

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty}\mathbb{E}\Big[\sum_{i\in\mathtt{T}_4^{(\ell)}}\mu_1^{(\ell)}\mu_i^{(\ell)}\Big]\\
=&\frac{1-\tilde{\lambda}_1^{2\Bbbk}}{1-\tilde{\lambda}_1^2}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{K}\tilde{\mathbf{e}}_1+\frac{1-\tilde{\lambda}_2^{2\Bbbk}}{1-\tilde{\lambda}_2^2}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{K}\tilde{\mathbf{e}}_2\big)\\
&+\sum_{i,j=3}^{(2m+1)^2}\frac{1-(\tilde{\lambda}_i\tilde{\lambda}_j)^{\Bbbk}}{1-\tilde{\lambda}_i\tilde{\lambda}_j}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\tilde{\lambda}_i\Big(\frac{1-(\tilde{\lambda}_2\tilde{\lambda}_i)^{\Bbbk}}{(1-\tilde{\lambda}_2\tilde{\lambda}_i)^2}-\frac{\Bbbk(\tilde{\lambda}_2\tilde{\lambda}_i)^{\Bbbk-1}}{1-\tilde{\lambda}_2\tilde{\lambda}_i}\Big)c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_2)\\
&+\tilde{\lambda}_1^{2\Bbbk}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{FB}_0\tilde{\mathbf{e}}_1+\tilde{\lambda}_2^{2\Bbbk}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{FB}_0\tilde{\mathbf{e}}_2\big)+\sum_{i,j=3}^{(2m+1)^2}\tilde{\lambda}_i^{\Bbbk}\tilde{\lambda}_j^{\Bbbk}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\Bbbk\tilde{\lambda}_2^{\Bbbk-1}\tilde{\lambda}_i^{\Bbbk}c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_2).
\end{aligned}
\tag{3.57}
$$

3.B.5 S^c for Iinfinite Number of Iterations and Finite Support Tree of Size $\Bbbk$

Size Correction

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty}x^{(\ell)}\mathbb{E}\big[(V^{\mathrm{G}_{\mathtt{T}}^{(\ell)}})'(1)\mu_1^{(\ell)}\big]\\
=&\lim_{\ell\to\infty}x^{(\ell)}\sum_d d\mathbb{E}\big[V_d^{\mathrm{G}_{\mathtt{T}}^{(\Bbbk)}}\mu_1\big]\\
=&\lim_{\ell\to\infty}\big((x^{(\ell)})^2\sum_d d\lambda_d\sum_{j=1}^{\Bbbk}\rho'(1)^j\lambda'(1)^{j-1}+ \qquad (3.58)\\
&+x^{(\ell)}\sum_{j=0}^{2\Bbbk}\mathbf{u}_m^T\prod_{h=\ell}^{\ell-j+1}\mathrm{V}^{(h)}\mathrm{C}^{(h-1)}\sum_d d\lambda_d\mathsf{a}^{(\ell-j)}(d)\big)\\
=&x^2\rho'(1)(1+\lambda'(1))\sum_{j=0}^{\Bbbk-1}\rho'(1)^j\lambda'(1)^j+x\mathbf{u}_m^T\sum_{j=0}^{2\Bbbk}(\mathrm{VC})^j\sum_d d\lambda_d\mathsf{a}(d)\\
=&x^2\rho'(1)(1+\lambda'(1))\frac{1-(\rho'(1)\lambda'(1))^{\Bbbk}}{1-\rho'(1)\lambda'(1)}+x\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk+1}}{1-\lambda_i}\hat{a}_i\mathbf{u}_m^T\mathbf{e}_i,
\end{aligned}
$$

where in the last step we define the vector $\hat{\mathbf{a}}=\sum_d d\lambda_d\mathbf{a}(d)$.

Degree Distribution Correction

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty} n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}[\tilde{\Delta}\mathsf{a}\mu_1]\big)\\
=&-\lim_{\ell\to\infty}\mathbf{u}_m^T\sum_{i=1}^{\ell}\prod_{h=\ell}^{i+1}\mathrm{V}^{(h)}\mathrm{C}^{(h-1)}\Big(\sum_d d\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{a}^{(i)}-\mathsf{a}^{(i)}(d))\\
&\qquad\qquad+\mathrm{V}^{(i)}\sum_d d\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{b}^{(i)}-\mathsf{b}^{(i)}(d))\Big)\\
\overset{l=\ell-i}{=}&-\lim_{\ell\to\infty}\mathbf{u}_m^T\sum_{l=0}^{\ell-1}\prod_{h=\ell}^{\ell-l+1}\mathrm{V}^{(h)}\mathrm{C}^{(h-1)}\Big(\sum_d d\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{a}^{(\ell-l)}-\mathsf{a}^{(\ell-l)}(d))\\
&\qquad\qquad+\mathrm{V}^{(\ell-l)}\sum_d d\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{b}^{(\ell-l)}-\mathsf{b}^{(\ell-l)}(d))\Big)\\
=&-\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\sum_d d\mathbb{E}[\mu_1V_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{a}-\mathsf{a}(d))+\mathrm{V}\sum_d d\mathbb{E}[\mu_1C_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}](\mathsf{b}-\mathsf{b}(d))\Big),
\end{aligned}
$$

where

$$
\begin{aligned}
\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}V_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}]=&\lim_{\ell\to\infty}x^{(\ell)}\sum_{k=1}^{\Bbbk}\rho'(1)^k\lambda'(1)^{k-1}\lambda_d\\
&+\lim_{\ell\to\infty}\sum_{k=0}^{2\Bbbk}\lambda_d\mathbf{u}_m^T\prod_{l=\ell}^{\ell-k+1}\mathrm{V}^{(l)}\mathrm{C}^{(l-1)}\mathsf{a}^{(\ell-k)}(d)\\
=&x\rho'(1)\sum_{k=0}^{\Bbbk-1}(\rho'(1)\lambda'(1))^k\lambda_d\\
&+\sum_{k=0}^{2\Bbbk}\lambda_d\mathbf{u}_m^T(\mathrm{VC})^k\mathsf{a}(d).
\end{aligned}
$$

Note that in the first equality, we already take into account the fact that $\ell > 2\Bbbk$. We also have

$$
\begin{aligned}
\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}C_d^{\mathrm{G}_{\mathrm{T}}^{(\Bbbk)}}]=&\lim_{\ell\to\infty}x^{(\ell)}\sum_{k=0}^{\Bbbk-1}(\rho'(1)\lambda'(1))^k\rho_d\\
&+\lim_{\ell\to\infty}\sum_{k=0}^{2\Bbbk-1}\rho_d\mathbf{u}_m^T\prod_{l=\ell}^{\ell-k+1}\mathrm{V}^{(l)}\mathrm{C}^{(l-1)}\mathrm{V}^{(\ell-k+1)}\mathsf{b}^{(\ell-k+1)}(d)\\
=&x\sum_{k=0}^{\Bbbk-1}(\rho'(1)\lambda'(1))^k\rho_d\\
&+\sum_{k=0}^{2\Bbbk-1}\rho_d\mathbf{u}_m^T(\mathrm{VC})^k\mathrm{V}\mathsf{b}(d).
\end{aligned}
$$

We finally have

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty} n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}[\tilde{\Delta}\mathsf{a}\mu_1]\big)\\
&= -\,\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\sum_d d\mathbb{E}[\mu_1 V_d^{\mathrm{G}_{\mathrm{T}}^{(\mathbb{k})}}](\mathsf{a}-\mathsf{a}(d)) + \mathrm{V}\sum_d d\mathbb{E}[\mu_1 C_d^{\mathrm{G}_{\mathrm{T}}^{(\mathbb{k})}}](\mathsf{b}-\mathsf{b}(d))\Big)\\
&= -\,\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Bigg(\sum_d d\lambda_d\Big(x\rho'(1)\sum_{k=0}^{\mathbb{k}-1}(\rho'(1)\lambda'(1))^k + \mathbf{u}_m^T\sum_{k=0}^{2\mathbb{k}}(\mathrm{VC})^k\mathsf{a}(d)\Big)(\mathsf{a}-\mathsf{a}(d))\\
&\quad + \mathrm{V}\sum_d d\rho_d\Big(x\sum_{k=0}^{\mathbb{k}-1}(\rho'(1)\lambda'(1))^k + \mathbf{u}_m^T\sum_{k=0}^{2\mathbb{k}-1}(\mathrm{VC})^k\mathrm{V}\mathsf{b}(d)\Big)(\mathsf{b}-\mathsf{b}(d))\Bigg)\\
&= -\,x\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\sum_d d\lambda_d\rho'(1)(\mathsf{a}-\mathsf{a}(d)) + \mathrm{V}\sum_d d\rho_d(\mathsf{b}-\mathsf{b}(d))\Big)\\
&\quad -\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\mathbf{u}_m^T\sum_d\sum_{k=0}^{2\mathbb{k}}(\mathrm{VC})^k d\lambda_d\mathsf{a}(d)(\mathsf{a}-\mathsf{a}(d))\\
&\qquad\qquad + \mathrm{V}\mathbf{u}_m^T\sum_d\sum_{k=0}^{2\mathbb{k}-1}(\mathrm{VC})^k d\rho_d\mathrm{V}\mathsf{b}(d)(\mathsf{b}-\mathsf{b}(d))\Big). \qquad (3.59)
\end{aligned}
$$

Let us look at the first part of (3.59). We have an expression of the form $\sum_{l\geq 0}(\mathrm{VC})^l\Delta\mathbf{c}$, where $\Delta\mathbf{c}$ is a vector of weight 0. The first eigenvalue of VC is equal to $\lambda'(1)\rho'(1)$, but it turns out that the first coefficient of the expansion of a vector of weight 0 in the basis formed by the eigen vectors of VC is equal to 0. We can thus write the firs part of (3.59) as

$$
-x\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}\mathbf{u}_m^T(\mathrm{I}-\mathrm{VC})^{-1}\Big(\sum_d d\lambda_d\rho'(1)(\mathsf{a}-\mathsf{a}(d)) + \mathrm{V}\sum_d d\rho_d(\mathsf{b}-\mathsf{b}(d))\Big), \qquad (3.60)
$$

even though that the first eigenvalue of VC is larger than 1.

Let us expand...

$$
\begin{aligned}
&-\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\mathbf{u}_m^T\sum_d\sum_{k=0}^{2\Bbbk}(\mathrm{VC})^k d\lambda_d\mathsf{a}(d)(\mathsf{a}-\mathsf{a}(d))\\
&\qquad\qquad +\mathrm{V}\mathbf{u}_m^T\sum_d\sum_{k=0}^{2\Bbbk-1}(\mathrm{VC})^k d\rho_d\mathrm{V}\mathsf{b}(d)(\mathsf{b}-\mathsf{b}(d))\Big)\\
&=-\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\mathbf{u}_m^T\sum_d\sum_{i=1}^{2m+1}\sum_{k=0}^{2\Bbbk}\lambda_i^k\tilde{a}_i(d)\mathbf{e}_i(\mathsf{a}-\mathsf{a}(d))\\
&\qquad\qquad +\mathrm{V}\mathbf{u}_m^T\sum_d\sum_{i=1}^{2m+1}\sum_{k=0}^{2\Bbbk-1}\lambda_i^k\tilde{b}_i(d)\mathbf{e}_i(\mathsf{b}-\mathsf{b}(d))\Big)\\
&=-\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk+1}}{1-\lambda_i}\sum_d\tilde{a}_i(d)\mathbf{u}_m^T\mathbf{e}_i(\mathsf{a}-\mathsf{a}(d))\\
&\qquad\qquad +\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk}}{1-\lambda_i}\sum_d\tilde{b}_i(d)\mathbf{u}_m^T\mathbf{e}_i\mathrm{V}(\mathsf{b}-\mathsf{b}(d))\Big).
\end{aligned}
\tag{3.61}
$$

Let us define the two vectors $\bar{\mathsf{a}}_i=\sum_d\tilde{a}_i(d)\mathbf{u}_m^T\mathbf{e}_i(\mathsf{a}-\mathsf{a}(d))$ and $\bar{\mathsf{b}}_i=\sum_d\tilde{b}_i(d)\mathbf{u}_m^T\mathbf{e}_i\mathrm{V}(\mathsf{b}-\mathsf{b}(d))$. We can then rewrite (3.61) as

$$
\begin{aligned}
&-\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\Big(\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk+1}}{1-\lambda_i}\sum_d\tilde{a}_i(d)\mathbf{u}_m^T\mathbf{e}_i(\mathsf{a}-\mathsf{a}(d))\\
&\qquad\qquad +\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk}}{1-\lambda_i}\sum_d\tilde{b}_i(d)\mathbf{u}_m^T\mathbf{e}_i\mathrm{V}(\mathsf{b}-\mathsf{b}(d))\Big)\\
&=-\Big(\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk+1}}{1-\lambda_i}\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\bar{\mathsf{a}}_i+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk}}{1-\lambda_i}\mathbf{u}_m^T\sum_{l\geq 0}(\mathrm{VC})^l\bar{\mathsf{b}}_i\Big)\\
&=-\Big(\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk+1}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{a}_{i,j}\mathbf{u}_m^T\mathbf{e}_j+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\Bbbk}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{b}_{i,j}\mathbf{u}_m^T\mathbf{e}_j\Big).
\end{aligned}
\tag{3.62}
$$

Note that in the last step the sum over j starts at 2. This can be done because as we said above, the first coefficient of the expansion of a vector of weight 0 in the basis formed by the eigen vectors of VC is equal to 0.

Messages Correction

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty}n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}[\hat{\Delta}\mathsf{a}\mu_1]\big)\\
&=\lim_{\ell\to\infty}\mathbf{u}_m^T\sum_{j=1}^{\ell-1}\prod_{h=\ell}^{\ell-j+1}\mathrm{V}^{(h)}\mathrm{C}^{(h-1)}\big(\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\Bbbk}|+|\mathsf{B}_{4,2\Bbbk}|)]\mathsf{a}^{(\ell-j)}-\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\Bbbk}|+|\mathsf{B}_{4,2\Bbbk}|)\mathsf{a}_\star^{(\ell-j)}]\big),
\end{aligned}
$$

where

$$
\begin{aligned}
\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\mathbb{k}}|+|\mathsf{B}_{4,2\mathbb{k}}|)] =& \lim_{\ell\to\infty}(\lambda'(1)\rho'(1))^{\mathbb{k}}x^{(\ell)} \\
&+\lim_{\ell\to\infty}\mathbf{u}_m^T\prod_{l=\ell}^{\ell-2\mathbb{k}+1}\mathrm{V}^{(l)}\mathrm{C}^{(l-1)}\sum_d(d-1)\lambda_d\mathsf{a}^{(\ell-2\mathbb{k})}(d) \\
=&(\lambda'(1)\rho'(1))^{\mathbb{k}}x+\mathbf{u}_m^T(\mathrm{VC})^{2\mathbb{k}}\sum_d(d-1)\lambda_d\mathsf{a}(d) \\
=&(\lambda'(1)\rho'(1))^{\mathbb{k}}x+\mathbf{u}_m^T(\mathrm{VC})^{2\mathbb{k}}\mathrm{V}\mathsf{b},
\end{aligned}
$$

Let us look at $\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}(|\mathsf{B}_{1,\mathbb{k}}|+|\mathsf{B}_{4,2\mathbb{k}}|)\mathsf{a}_\star^{(\ell-j)}] = \lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{1,\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}]+\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}]$. We have

$$
\begin{aligned}
&\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{1,\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}] \\
=&\lim_{\ell\to\infty}\sum_{r=-m}^{m}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{1,\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}\mid\nu_1^{(\ell-j-\mathbb{k})}=r]\mathbb{P}\{\nu_1^{(\ell-j-\mathbb{k})}=r\} \\
=&\lim_{\ell\to\infty}\sum_{r=-m}^{m}\mathbb{E}[\mu_1^{(\ell)}\mid\nu_1^{(\ell-j-\mathbb{k})}=r]\mathbb{E}[|\mathsf{B}_{1,\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}\mid\nu_1^{(\ell-j-\mathbb{k})}=r]\mathbb{P}\{\nu_1^{(\ell-j-\mathbb{k})}=r\} \\
=&\lim_{\ell\to\infty}\sum_{r=-m}^{m}\mathbb{P}\{\nu_1^{(\ell)}=m,\nu_1^{(\ell-j-\mathbb{k})}=r\}\mathbb{E}[|\mathsf{B}_{1,\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}\mid\nu_1^{(\ell-j-\mathbb{k})}=r] \\
=&\sum_{r=-m}^{m}\mathrm{P}_{m,r}(\mathrm{VC})^{\mathbb{k}}\mathbf{e}_r \\
=&(\mathrm{VC})^{\mathbb{k}}\mathrm{P}_m
\end{aligned}
$$

where P is the variable flipping matrix at fixed point and P_m is its m^{th} row. Then

$$
\begin{aligned}
&\lim_{\ell\to\infty}\mathbb{E}[\mu_1^{(\ell)}|\mathsf{B}_{4,2\mathbb{k}}|\mathsf{a}_\star^{(\ell-j)}] \\
=&\lim_{\ell\to\infty}\lim_{n\to\infty}\sum_{r=-m}^{m}\Big(\prod_{j=0}^{\mathbb{k}-1}\hat{\mathrm{B}}^{2\mathbb{k},j}\mathrm{B}^{2\mathbb{k},j+1}\mathbf{c}_{\text{init}}(m)\Big)^T\mathrm{F}\mathrm{B}^{2\mathbb{k},0}\prod_{j=0}^{\mathbb{k}-1}\hat{\mathrm{B}}^{2\mathbb{k},j}\mathrm{B}^{2\mathbb{k},j+1}\mathbf{c}_{\text{init}}(r)\mathbf{u}_r \\
=&\sum_{r=-m}^{m}\Big((\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\mathbf{c}_{\text{init}}(m)\Big)^T\mathrm{F}\mathrm{B}_0(\hat{\mathrm{B}}\mathrm{B})^{\mathbb{k}}\mathbf{c}_{\text{init}}(r)\mathbf{u}_r
\end{aligned}
$$

where

$$
\begin{aligned}
\lim_{\ell\to\infty}\big(\mathbf{c}_{\text{init}}(i)\big)_{(s,r)} &= \mathbb{P}\{\nu_{\hat{0}\leftarrow 0}^{(\ell)}=i,\nu_{\hat{0}\to 0}^{(\ell-\mathbb{k})}=s\mid\nu_{\hat{0}\leftarrow 0}^{(\ell)}=r\} \\
&= \lim_{\ell\to\infty}\mathbb{P}\{\nu_{\hat{0}\to 0}^{(\ell-\mathbb{k})}=s\}\mathbb{1}_{\{i=r\}} \\
&= \lim_{\ell\to\infty}\mathbf{e}_s^T\mathsf{b}^{\ell-\mathbb{k}}\mathbb{1}_{\{i=r\}} \\
&= \mathbf{e}_s^T\mathsf{b}\mathbb{1}_{\{i=r\}}.
\end{aligned}
$$

Finally we have

$$
\begin{aligned}
&\lim_{\ell\to\infty}\lim_{n\to\infty} n\Lambda'(1)\mathfrak{E}\big(\mathbb{E}[\hat{\Delta}\mathsf{a}\mu_1]\big)\\
=&\mathbf{u}_m^T\sum_{j\geq 1}(\mathrm{VC})^j\Bigg(\Big((\lambda'(1)\rho'(1))^{\Bbbk}x\mathsf{a}-(\mathrm{VC})^{\Bbbk}\mathrm{P}_m\Big)\\
&+\Big(\mathbf{u}_m^T(\mathrm{VC})^{2\Bbbk}\mathrm{V}\mathsf{ba}-\sum_{r=-m}^{m}\Big((\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c}_{\mathrm{init}}(m)\Big)^T\mathrm{FB}_0(\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c}_{\mathrm{init}}(r)\mathbf{u}_r\Big)\Bigg)\\
=&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\Bigg(\Big((\lambda'(1)\rho'(1))^{\Bbbk}x\mathsf{a}-(\mathrm{VC})^{\Bbbk}\mathrm{P}_m\Big)\\
&+\Big(\sum_{i=1}^{2m+1}\lambda_i^{2\Bbbk}v_i\mathbf{u}_m^T\mathbf{e}_i\mathsf{a}-\sum_{r=-m}^{m}\Big((\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c}_{\mathrm{init}}(m)\Big)^T\mathrm{FB}_0(\hat{\mathrm{B}}\mathrm{B})^{\Bbbk}\mathbf{c}_{\mathrm{init}}(r)\mathbf{u}_r\Big)\Bigg),
\end{aligned}
$$

where in the last step we again use the fact that the first coefficient of the expansion of a vector of weight 0 in the basis formed by the eigen vectors of VC is equal to 0.

3.B.6 Put It Together

$$
\begin{aligned}
&\mathtt{T}_1+\mathtt{T}_2+\mathtt{T}_3+\mathtt{T}_4+\mathrm{ddc}+\mathrm{mc}+\mathrm{sc}\\
=&\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\sum_{i=1}^{2m+1}\lambda_i^{\mathbb{k}}p_i\mathbf{e}_i\\
&+x^2\rho'(1)\frac{1-(\lambda'(1)\rho'(1))^{\mathbb{k}}}{1-\lambda'(1)\rho'(1)}\\
&+\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\sum_{i=1}^{2m+1}\lambda_i^{2\mathbb{k}}p_i\mathbf{e}_i\\
&+\frac{1-\tilde{\lambda}_1^{2\mathbb{k}}}{1-\tilde{\lambda}_1^2}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{K}\tilde{\mathbf{e}}_1+\frac{1-\tilde{\lambda}_2^{2\mathbb{k}}}{1-\tilde{\lambda}_2^2}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{K}\tilde{\mathbf{e}}_2\big)\\
&+\sum_{i,j=3}^{(2m+1)^2}\frac{1-(\tilde{\lambda}_i\tilde{\lambda}_j)^{\mathbb{k}}}{1-\tilde{\lambda}_i\tilde{\lambda}_j}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\tilde{\lambda}_i\Big(\frac{1-(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}}}{(1-\tilde{\lambda}_2\tilde{\lambda}_i)^2}-\frac{\mathbb{k}(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}-1}}{1-\tilde{\lambda}_2\tilde{\lambda}_i}\Big)c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_2)\\
&+\tilde{\lambda}_1^{2\mathbb{k}}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{FB}_0\tilde{\mathbf{e}}_1+\tilde{\lambda}_2^{2\mathbb{k}}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{FB}_0\tilde{\mathbf{e}}_2\big)+\sum_{i,j=3}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}\tilde{\lambda}_j^{\mathbb{k}}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}\tilde{\lambda}_i^{\mathbb{k}}c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_2)\\
&+x\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}\mathbf{u}_m^T(\mathrm{I}-\mathrm{VC})^{-1}\Big(\sum_d d\lambda_d\rho'(1)(\mathsf{a}-\mathsf{a}(d))+\mathrm{V}\sum_d d\rho_d(\mathsf{b}-\mathsf{b}(d))\Big)\\
&+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}+1}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{a}_{i,j}\mathbf{u}_m^T\mathbf{e}_j+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{b}_{i,j}\mathbf{u}_m^T\mathbf{e}_j\\
&-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\Big((\lambda'(1)\rho'(1))^{\mathbb{k}}x\mathsf{a}-\sum_{i=1}^{2m+1}\lambda_i^{\mathbb{k}}p_i\mathbf{e}_i\Big)\\
&-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\left(\sum_{i=1}^{2m+1}\lambda_i^{2\mathbb{k}}v_i\mathbf{u}_m^T\mathbf{e}_i\mathsf{a}\right.\\
&\quad\left.-\sum_{r=-m}^{m}\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}c_i\tilde{\mathbf{e}}_i+\mathbb{k}\lambda_2^{\mathbb{k}-1}c_3\tilde{\mathbf{e}}_2\Big)^T\mathrm{FB}_0\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}c_i(r)\tilde{\mathbf{e}}_i+\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}c_3(r)\tilde{\mathbf{e}}_2\Big)\mathbf{u}_r\right)\\
&-x^2\rho'(1)(1+\lambda'(1))\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}-x\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}+1}}{1-\lambda_i}\hat{a}_i\mathbf{u}_m^T\mathbf{e}_i.
\end{aligned}
\tag{3.63}
$$

$$
\begin{aligned}
&\mathtt{T}_1+\mathtt{T}_2+\mathtt{T}_3+\mathtt{T}_4+\mathrm{ddc}+\mathrm{mc}+\mathrm{sc}\\
=&2\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathrm{P}_m-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\sum_{i=1}^{2m+1}\lambda_i^{2\mathbb{k}}p_i\mathbf{e}_i\\
&+\frac{1-\tilde{\lambda}_1^{2\mathbb{k}}}{1-\tilde{\lambda}_1^2}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{K}\tilde{\mathbf{e}}_1+\frac{1-\tilde{\lambda}_2^{2\mathbb{k}}}{1-\tilde{\lambda}_2^2}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{K}\tilde{\mathbf{e}}_2\big)\\
&+\sum_{i,j=3}^{(2m+1)^2}\frac{1-(\tilde{\lambda}_i\tilde{\lambda}_j)^{\mathbb{k}}}{1-\tilde{\lambda}_i\tilde{\lambda}_j}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\tilde{\lambda}_i\Big(\frac{1-(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}}}{(1-\tilde{\lambda}_2\tilde{\lambda}_i)^2}-\frac{\mathbb{k}(\tilde{\lambda}_2\tilde{\lambda}_i)^{\mathbb{k}-1}}{1-\tilde{\lambda}_2\tilde{\lambda}_i}\Big)c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{K}\tilde{\mathbf{e}}_2)\\
&+\tilde{\lambda}_1^{2\mathbb{k}}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{FB}_0\tilde{\mathbf{e}}_1+\tilde{\lambda}_2^{2\mathbb{k}}c_2c_3\big(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_3+\tilde{\mathbf{e}}_3^T\mathrm{FB}_0\tilde{\mathbf{e}}_2\big)+\sum_{i,j=3}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}\tilde{\lambda}_j^{\mathbb{k}}c_ic_j\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_j\\
&+\sum_{i=3}^{(2m+1)^2}\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}\tilde{\lambda}_i^{\mathbb{k}}c_3c_i(\tilde{\mathbf{e}}_2^T\mathrm{FB}_0\tilde{\mathbf{e}}_i+\tilde{\mathbf{e}}_i^T\mathrm{FB}_0\tilde{\mathbf{e}}_2)\\
&+x\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}\mathbf{u}_m^T(\mathrm{I}-\mathrm{VC})^{-1}\Big(\sum_d d\lambda_d\rho'(1)(\mathsf{a}-\mathsf{a}(d))+\mathrm{V}\sum_d d\rho_d(\mathsf{b}-\mathsf{b}(d))\Big)\\
&+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}+1}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{a}_{i,j}\mathbf{u}_m^T\mathbf{e}_j+\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}}}{1-\lambda_i}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{b}_{i,j}\mathbf{u}_m^T\mathbf{e}_j\\
&-(\lambda'(1)\rho'(1))^{\mathbb{k}}x\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathsf{a}\\
&-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\left(\sum_{i=1}^{2m+1}\lambda_i^{2\mathbb{k}}v_i\mathbf{u}_m^T\mathbf{e}_i\mathsf{a}\right.\\
&\qquad\left.-\sum_{r=-m}^{m}\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}c_i\tilde{\mathbf{e}}_i+\mathbb{k}\lambda_2^{\mathbb{k}-1}c_3\tilde{\mathbf{e}}_2\Big)^T\mathrm{FB}_0\Big(\sum_{i=1}^{(2m+1)^2}\tilde{\lambda}_i^{\mathbb{k}}c_i(r)\tilde{\mathbf{e}}_i+\mathbb{k}\tilde{\lambda}_2^{\mathbb{k}-1}c_3(r)\tilde{\mathbf{e}}_2\Big)\mathbf{u}_r\right)\\
&-x^2\rho'(1)\lambda'(1)\frac{1-(\rho'(1)\lambda'(1))^{\mathbb{k}}}{1-\rho'(1)\lambda'(1)}-x\sum_{i=1}^{2m+1}\frac{1-\lambda_i^{2\mathbb{k}+1}}{1-\lambda_i}\hat{a}_i\mathbf{u}_m^T\mathbf{e}_i. \qquad (3.64)
\end{aligned}
$$

We know that the terms wich multiplies $\lambda_1^{\mathbb{k}}$ and $\lambda_1^{2\mathbb{k}}$ should be equal to 0. According to the derivations above, the term which multiplies $\lambda_1^{\mathbb{k}}$, i.e., $(\rho'(1)\lambda'(1))^{\mathbb{k}}$ is equal to:
$\frac{-x}{1-\rho'(1)\lambda'(1)}\mathbf{u}_m^T(\mathrm{I}-\mathrm{VC})^{-1}\Big(\sum_d d\lambda_d\rho'(1)(\mathsf{a}-\mathsf{a}(d))+\mathrm{V}\sum_d d\rho_d(\mathsf{b}-\mathsf{b}(d))\Big)-x\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\mathsf{a}+x^2\frac{\rho'(1)\lambda'(1)}{1-\rho'(1)\lambda'(1)}$,
and the one which multiplies $\lambda_1^{2\mathbb{k}}$ is equal to:
$-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}p_1\mathbf{e}_1-\frac{1}{1-\tilde{\lambda}_1^2}c_1^2\tilde{\mathbf{e}}_1^T\mathrm{K}\tilde{\mathbf{e}}_1+c_1^2\tilde{\mathbf{e}}_1^T\mathrm{FB}_0\tilde{\mathbf{e}}_1-\frac{\lambda_1}{1-\lambda_1}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{a}_{1,j}\mathbf{u}_m^T\mathbf{e}_j-$

$\frac{1}{1-\lambda_1}\sum_{j=2}^{2m+1}\frac{1}{1-\lambda_j}\bar{b}_{1,j}\mathbf{u}_m^T\mathbf{e}_j-\mathbf{u}_m^T\mathrm{VC}(\mathrm{I}-\mathrm{VC})^{-1}\bigg(v_1\mathbf{u}_m^T\mathbf{e}_1\mathsf{a}-c_1\sum_{r=-m}^{m}c_1(r)\tilde{\mathbf{e}}_1^T\mathrm{FB}_0\tilde{\mathbf{e}}_1\bigg)+$
$x\frac{\lambda_1}{1-\lambda_1}\hat{a}_1\mathbf{u}_m^T\mathbf{e}_1.$
One can check that for any particular example, these two terms are indeed equal to 0. According to this observation, we are now able to let the size of $\mathrm{G}_{\mathtt{T}}^{(\ell)}$ going to infinity, which means taking the limit $k \to \infty$. To that end, we will use the fact that for $i \geq 2$, $\lambda_i < 1$, when $x > x^{\mathrm{MP}}$. We can thus write

$$\begin{aligned}&\lim_{\Bbbk\to\infty}(\mathtt{T}_1+\mathtt{T}_2+\mathtt{T}_3+\mathtt{T}_4+\mathrm{ddc}+\mathrm{mc}+\mathrm{sc})\\ =&\frac{2\lambda_2}{(1-\lambda_2^2)^2}c_3^2\tilde{\mathbf{e}}_2^T\mathrm{K}\tilde{\mathbf{e}}_3+\frac{1}{(1-\lambda_2)^2}\big(\bar{a}_{2,2}\mathbf{u}_m^T\mathbf{e}_2+\bar{b}_{2,2}\mathbf{u}_m^T\mathbf{e}_2\big)+O\bigg(\frac{1}{1-\lambda_2}\bigg).\end{aligned} \tag{3.65}$$

4 Flipping

4.1 Introduction

In this chapter we look more in depth at the “flipping” phenomenon. As briefly discussed earlier, this refer to the following. Assume we are transmitting above the threshold. Consider an infinite tree and consider density evolution. Then generically for a well-defined message- passing algorithm the message *densities* converge but the actual individual messages do not – they flip. Although this flipping can be observed routinely when decoding sparse graph codes, it is usually attributed to the cycles which are present in a finite-length code. We show that in fact this phenomenon is also present on a tree. Besides being of independent interest when trying to understand sparse graph codes, the flipping properties are crucial in deriving the scaling law. There is one prominent exception to the above rule. For the BP decoding we will show that even above threshold the messages themselves converge.

In the sequel we assume a message passing algorithm over an alphabet $\mathcal{M}$. This alphabet can either be discrete. In this case we assume that it is of the form $\mathcal{M} = \{-m, \cdots, m\}$, where $m \in \mathbb{N}$. Or it can be continuous. Our notation will reflect the case of a discrete alphabet.

4.2 Stability Condition for Flipping Matrix

We assume an infinite tree and an infinite number of iterations so that all densities have converged. We also think of transmitting above the threshold so that we get a not-trivial fixed point of density evolution.

Let p and q denote the stationary flipping probabilities. More precisely, we pick two time instances which are fixed non-zero time apart (time here is measured by an integer since it refers to an iteration). We will see that the actual value of the time separation does not play a role. Since further we assumed that the system has already converged to its steady state, the actual time instance does not play a role either. Hence we will not specify either of them further.

Then $p_{i,j}$ is the joint probability that at the first time instance the message emitted by a given variable node along a given edge is i and that at the later time instance this message is j. Some thought shows that the flipping probabilities must fulfill the following fixed point equations:

$$\begin{aligned} q &= p^{\boxast(\mathtt{r}-1)}, \\ p &= V(c \star q^{\star(\mathtt{l}-1)}), \end{aligned}$$

To be more precise. Assume that the message-passing decoder fulfills the following conditions. At the check nodes the message passing rule can be specified pairwise. This means that if we have a node of degree $\mathtt{r}$ then the output can be computed by performing $\mathtt{r}-2$ pair-wise operations. At the variable nodes we assume that the sum of all incoming messages (including the received one) is a sufficient statistic. If these conditions are fulfilled then the following computations are made somewhat simpler, but they are not absolutely necessary. With this restriction let us first look at the check node side. It suffices to define the pairwise convolution of two flipping matrices at the check node side. For two input flipping matrices $p^{(1)}$ and $p^{(2)}$ we then have

$$(p^{(1)} \boxast p^{(2)})_{i,j} = \sum_{i_1,i_2,j_1,j_2} p^{(1)}_{i_1,j_1} p^{(1)}_{i_2,j_2} \mathbb{1}_{\{c(i_1,i_2)=i\}} \mathbb{1}_{\{c(j_1,j_2)=j\}}.$$

This defines $p^{\boxast(\mathtt{r}-1)}$ and, hence, q.

At the variable nodes $c \star q^{\star(\mathtt{l}-1)}$ has the following meaning. Let $q^{(1)}$ and $q^{(2)}$ denote two flipping matrices, where the matrices might be over an extended alphabet, i.e., the matrices are no longer necessarily over $\mathcal{M}$. It is easiest to assume that the matrices are infinite with entries in $\mathbb{Z}$ but that only a finite submatrix has non-zero entries. We define

$$(q^{(1)} \boxast q^{(2)})_{i,j} = \sum_{i_1,i_2,j_1,j_2} q^{(1)}_{i_1,j_1} q^{(1)}_{i_2,j_2} \mathbb{1}_{\{i_1+i_2=i\}} \mathbb{1}_{\{j_1+j_2=j\}}.$$

Finally, the projection operator V maps a possibly extended matrix back onto the alphabet $\mathcal{M}$ by mapping any element outside the range $\{-m, \cdots, m\}$ back to either the element $-m$ or m.

Note that the flipping probabilities must have the following marginals. We must have

$$\begin{aligned} \sum_i p_{i,j} &= \sum_i p_{j,i} = x_j, \\ \sum_i q_{i,j} &= \sum_i q_{j,i} = y_j. \end{aligned}$$

If there is no flipping then the flipping matrices must therefore be the following. Let $\bar{p}$ and $\bar{q}$ denote the following “stable” flipping probabilities. We have

$$\begin{aligned}\bar{p}_{i,j} &= x_i \delta_{i,j},\\ \bar{q}_{i,j} &= y_i \delta_{i,j},\end{aligned}$$

where x and y are the fixed point densities at the output of the variable and check node, respectively.

To show that there is flipping we will show that these “stable” matrices are not stable.

4.3 Gallager A

Let us look at the example of the Gallager algorithm A and a $(\mathtt{l}, \mathtt{r})$ ensemble. In this case a flipping matrix has the form

$$\begin{aligned}p &= \begin{pmatrix} a & x_{-1} - a \\ x_{-1} - a & x_1 - x_{-1} + a \end{pmatrix}\\ q &= \begin{pmatrix} a & y_{-1} - a \\ y_{-1} - a & y_1 - y_{-1} + a \end{pmatrix}\end{aligned}$$

The time invariant matrices have the from

$$\begin{aligned}\bar{p} &= \begin{pmatrix} x_{-1} & 0 \\ 0 & x_1 \end{pmatrix}\\ \bar{q} &= \begin{pmatrix} y_{-1} & 0 \\ 0 & y_1 \end{pmatrix}\end{aligned}$$

We will now show that these matrices are not stable. Assume we have a small perturbation.

$$p = \begin{pmatrix} x_{-1} + \Delta & -\Delta \\ -\Delta & x_1 + \Delta \end{pmatrix}$$

Consider first the check step. We think of Δ as very small. Then up to terms linear in Δ

$$\begin{pmatrix} y_{-1} + (\mathtt{r} - 1)\Delta & -(\mathtt{r} - 1)\Delta \\ -(\mathtt{r} - 1)\Delta & y_1 + (\mathtt{r} - 1)\Delta \end{pmatrix}$$

Now close the circle by considering the variable node step.

$$\begin{pmatrix} x_{-1} + \tilde{\Delta} & -\tilde{\Delta} \\ -\tilde{\Delta} & x_1 + (\mathtt{r} - 1)\tilde{\Delta} \end{pmatrix}$$

where $\tilde{\Delta} = (\mathtt{l}-1)(\mathtt{r}-1)(y_{-1}^{\mathtt{l}-2}(1-\epsilon) + y_1^{\mathtt{l}-2}\epsilon)\Delta$. Therefore if

$$(\mathtt{l}-1)(\mathtt{r}-1)(y_{-1}^{\mathtt{l}-2}(1-\epsilon) + y_1^{\mathtt{l}-2}\epsilon) > 1$$

then the time invariant matrix is not stable.

Consider e.g. the $(3,3)$ ensemble. In this case we have at the threshold $y_{-1} = 0.284307$ and $\epsilon^* = 0.2230463$. A quick check shows that the left hand side is equal to 1.5221 and that it is increasing for increasing channel parameters. We conclude that we should expect a non-trivial flipping matrix. Indeed this is what we observe.

We know from the above that the time-invariant matrix is a fixed point of the recursion but is not stable. In particular, if we start with $a = x_{-1} - \delta$, for δ sufficiently small, then we get a value of a strictly smaller than $x_{-1} - \delta$ after one recursion. On the other hand, the matrix with $a = 0$ is not a fixed point of density evolution and gives us a value strictly larger than 0 after one iteration. From this we conclude that the recursion has at least one stable fixed-point matrix. If we can show that it has exactly one then it follows that for any time difference the flipping probabilities converge to this fixed point matrix.

4.4 General Stability Condition For Time-Invariant Flipping Matrix

Consider the time-invariant flipping matrices

$$\bar{p}_{i,j} = x_i \delta_{i,j},$$
$$\bar{q}_{i,j} = y_i \delta_{i,j}.$$

To simplify matters assume for right now that x is component-wise strictly positive. Consider any small deviation from it. Represent this deviation as

$$p = \bar{p} + \sum_{-m \leq i < j \leq m} \Delta^{(i,j)} A^{(i,j)},$$

where $A^{(i,j)}$ is a $(2m+1) \times (2m+1)$ matrix defined by

$$A_{l,k}^{(i,j)} = \begin{cases} -1, & l = k = i, \\ -1, & l = k = j, \\ 1, & l = i, k = j \\ 1, & l = j, k = i, \\ 0, & \text{otherwise.} \end{cases}$$

Consider now the evolution of this deviation for one round. First consider the convolution at check nodes. Up to terms linear in the deviation we have

$$\begin{aligned} q &= p^{\boxast(\mathtt{r}-1)} \\ &= (\bar{p} + \sum_{-m\le i<j\le m} \Delta^{(i,j)} A^{(i,j)})^{\boxast(\mathtt{r}-1)} \\ &= \bar{p}^{\boxast(\mathtt{r}-1)} + (\mathtt{r}-1)\bar{p}^{\boxast(\mathtt{r}-2)} \boxast \sum_{-m\le i<j\le m} \Delta^{(i,j)} A^{(i,j)} \\ &= \bar{q} + (\mathtt{r}-1) \sum_{-m\le i<j\le m} \Delta^{(i,j)} \bar{p}^{\boxast(\mathtt{r}-2)} \boxast A^{(i,j)} \end{aligned}$$

Let us define $\tilde{A}^{(i,j)} = \bar{p}^{\boxast(\mathtt{r}-2)} \boxast A^{(i,j)}$. Then we have

$$\begin{aligned} \tilde{A}^{(i,j)}_{k,l} &= \sum_h x_h^{\boxast(\mathtt{r}-2)} \left(\mathbb{1}_{\{c(h,i)=k\}}\mathbb{1}_{\{c(h,j)=l\}} + \mathbb{1}_{\{c(h,j)=k\}}\mathbb{1}_{\{c(h,i)=l\}}\right) \\ &- \sum_h x_h^{\boxast(\mathtt{r}-2)} \left(\mathbb{1}_{\{c(h,i)=k\}}\mathbb{1}_{\{c(h,i)=l\}} + \mathbb{1}_{\{c(h,j)=k\}}\mathbb{1}_{\{c(h,j)=l\}}\right) \end{aligned}$$

and we can write

$$\tilde{A}^{(i,j)} = \sum_{-m\le l<k\le m} \tilde{A}^{(i,j)}_{k,l} A^{(k,l)}.$$

Then

$$\begin{aligned} q &= \bar{q} + (\mathtt{r}-1) \sum_{-m\le i<j\le m} \Delta^{(i,j)} \sum_{-m\le k<l\le m} \tilde{A}^{(i,j)}_{k,l} A^{(k,l)} \\ &= \bar{q} + \sum_{-m\le k<l\le m} (\mathtt{r}-1) \sum_{-m\le i<j\le m} \Delta^{(i,j)} \tilde{A}^{(i,j)}_{k,l} A^{(k,l)} \\ &= \bar{q} + \sum_{-m\le k<l\le m} \tilde{\Delta}^{(k,l)} A^{(k,l)} \end{aligned}$$

Consider the vector Δ of length $\binom{2m+1}{2}$ with components $\Delta_k = \Delta^{\phi(k)}$ where ϕ is a bijective map from $\{1,\dots,\binom{2m+1}{2}\}$ to a "valid" pair (i,j).

Then we have

$$\tilde{\Delta} = (\mathtt{r}-1)C\Delta$$

where $C_{i,j} = \tilde{A}^{\phi(j)}_{\phi(i)}$.

Now let us consider a variable step.

$$\begin{aligned} p &= V\left(c \star \left(\bar{q} + \sum_{-m\le k<l\le m} \tilde{\Delta}^{(k,l)} A^{(k,l)}\right)^{\star(\mathtt{l}-1)}\right) \\ &= V\left(c \star \left(\bar{q}^{\star(\mathtt{l}-1)} + (\mathtt{l}-1) \sum_{-m\le k<l\le m} \tilde{\Delta}^{(k,l)} \bar{q}^{\star(\mathtt{l}-2)} \star A^{(k,l)}\right)\right) \\ &= V\left(c \star \bar{q}^{\star(\mathtt{l}-1)}\right) \\ &+ \sum_{-m\le k<l\le m} (\mathtt{l}-1)\tilde{\Delta}^{(k,l)} V\left(c \star \bar{q}^{\star(\mathtt{l}-2)} \star A^{(k,l)}\right) \end{aligned}$$

Let us define $\hat{A}^{(k,l)} = V\left(c \star \bar{q}^{\star(\mathtt{l}-2)} \star A^{(k,l)}\right)$. Thus

$$\begin{aligned} p &= \bar{p} + \sum_{-m\le k<l\le m} (\mathtt{l}-1)\tilde{\Delta}^{(k,l)} \sum_{-m\le i<j\le m} \hat{A}^{(k,l)}_{i,j} A^{(i,j)} \\ &= \bar{p} + \sum_{-m\le i<j\le m} \sum_{-m\le k<l\le m} (\mathtt{l}-1)\tilde{\Delta}^{(k,l)} \hat{A}^{(k,l)}_{i,j} A^{(i,j)} \\ &= \bar{p} + \sum_{-m\le i<j\le m} \hat{\Delta}^{(i,j)} A^{(i,j)} \end{aligned}$$

As before, we have

$$\hat{\Delta} = (\mathtt{l}-1)V\tilde{\Delta}$$

where $V_{i,j} = \hat{A}^{\phi(j)}_{\phi(i)}$.

It follows that after one complete iteration the deviation coefficients have evolved according to the equation

$$\hat{\Delta} = \gamma V C \Delta.$$

Therefore, if γVC has a spectral radius strictly larger than 1 then this recursion is not stable and flipping occurs. But if the spectral radius is smaller than 1 then this fixed point is stable.

4.5 Time-Invariance of BP Decoder

Let us show that the messages of the BP converge on a tree. In the sequel it is convenient to assume that the messages which are passed along the tree are posteriors. Consider a tree of depth ℓ. We will then show that for most instances of channel realizations the posterior of the message sent along the root edge changes by at most δ_ℓ and that δ_ℓ tends to zero as ℓ tends to infinity.

For notational simplicity assume that the channel has a discrete output alphabet. The statement remains true however for the general case. Let X denote the value of the root bit, $X \in \{\pm 1\}$. Let Y_i, $i = 0, 1, \cdots$, denote the collection of observations of the tree channel at *level* i. Let Y_0^ℓ denote all observations up to level ℓ.

Lemma 4.1. [Convergence of Messages for the BP Decoder] Consider transmission over a tree channel. Then

$$\begin{aligned}&\sum_{x, y_0^\ell} p(y_0^\ell) |p(x \mid y_0^\ell) - p(x \mid y_0^{\ell-1})| \\ &\leq \sqrt{2}\sqrt{H(X \mid Y_0^\ell) - H(X \mid Y_0^{\ell-1})}.\end{aligned}$$

Further, if

$$\mathcal{G} = \{y_0^\ell : \|p(x \mid y_0^\ell) - p(x \mid y_0^{\ell-1})\|_1 \leq \rho\},$$

then

$$P\{Y_0^\ell \notin \mathcal{G}\} \leq \frac{\sqrt{2}}{\rho}\sqrt{H(X \mid Y_0^\ell) - H(X \mid Y_0^{\ell-1})}.$$

Discussion: Note that $H(X \mid Y_0^\ell)$ is decreasing as a function of ℓ. Further, it is upper bounded by 1 and lower bounded by 0. It follows that it converges to a limit. Therefore, $H(X \mid Y_0^{\ell-1}) - H(X \mid Y_0^{\ell 1})$ converges to 0 as a function of ℓ. The statement therefore says that for most instances of the observations the messages change only slightly as we add one more layer of observations.

Proof. We have

$$\begin{aligned}&\sum_{x, y_0^\ell} p(y_0^\ell) |p(x \mid y_0^\ell) - p(x \mid y_0^{\ell-1})| \\ =&\sum_{x, y_0^\ell} |p(x, y_0^\ell) - p(x \mid y_0^{\ell-1}) p(y_0^\ell)| \\ \leq&\sqrt{2D(p(x, y_0^\ell) \| p(x \mid y_0^{\ell-1}) p(y_0^\ell))} \\ =&\sqrt{2}\sqrt{H(X \mid Y_0^\ell) - H(X \mid Y_0^{\ell-1})},\end{aligned}$$

where the second step is the so-called "Pinsker inequality."

□

4.6 Flipping after a Finite Number of Iterations

In the previous sections we considered the flipping probabilities once the system has reached steady state. But it is also of interest to compute the flipping probabilities in the transient phase before the decoder has converged. We will see that in this case the flipping probabilities can be expressed by means of a recursion.

Consider $\nu^{(j)}$ to be a variable-to-check message in iteration j. Let us show how to compute $\mathbb{P}\{\nu^{(\ell)} = r, \nu^{(\ell-k)} = s\}$, where r and s are two message values. We call this the ”flipping” probability recursion.

Consider a variable-to-check message ν and let us define the $2m+1 \times 2m+1$ dimensional variable-to-check flipping matrix as $\mathrm{P}^{(t_1,t_2)}_{i,j} = \mathbb{P}\{\nu^{(t_1)} = i, \nu^{(t_2)} = j\}$. Note that the indices i and j go from $-m$ to m. In the same way considering a check-to-variable message $\hat{\nu}$, we define the $2m+1 \times 2m+1$ dimensional check-to-variable flipping matrix as $\mathrm{Q}^{(t_1,t_2)}_{i,j} = \mathbb{P}\{\hat{\nu}^{(t_1)} = i, \hat{\nu}^{(t_2)} = j\}$. Finally we define the $2m+1 \times 2m+1$ dimensional “channel” matrix as $(\mathrm{C}_{\mathsf{BMS}})_{i,j} = \mathbb{1}_{\{i=j\}}(\mathsf{a}_{\mathsf{BMS}})_i$, where $\mathsf{a}_{\mathsf{BMS}}$ is the density coming from the channel. Consider the matrix K of size $(2k+1)\times(2k+1)$ and the matrix L of size $(2l+1)\times(2l+1)$. We define the operator $\odot : \mathbb{R}^{(2k+1)\times(2k+1)} \times \mathbb{R}^{(2l+1)\times(2l+1)} \mapsto \mathbb{R}^{(2(k+l)+1)\times(2(k+l)+1)}$ such that

$$(\mathrm{K} \odot \mathrm{L})_{i,j} = \sum_{r,s\in\{-k,\dots,k\}} \sum_{u,v\in\{-l,\dots,l\}} \mathrm{K}_{r,s}\mathrm{L}_{u,v}\mathbb{1}_{\{r+u=i\}}\mathbb{1}_{\{s+v=j\}}.$$

Consider two matrices M and N of size $2m+1 \times 2m+1$, we also define the operator $\boxdot : \mathbb{R}^{(2m+1)\times(2m+1)} \times \mathbb{R}^{(2m+1)\times(2m+1)} \mapsto \mathbb{R}^{(2m+1)\times(2m+1)}$ such that

$$(\mathrm{M} \boxdot \mathrm{N})_{i,j} = \sum_{r,s,u,v\in\{-m,\dots,m\}} \mathrm{M}_{r,s}\mathrm{N}_{u,v}\mathbb{1}_{\{\mathfrak{Q}(r\boxplus u)=i\}}\mathbb{1}_{\{\mathfrak{Q}(s\boxplus v)=j\}},$$

where $\boxplus$ is the operator which corresponds to the check node rule and $\mathfrak{Q}$ is the operation which performs the quantization as defined in Section 3.3. Finally, by an abuse of notation, we extend the definition of the operator $\mathfrak{Q}$, to the framework of flipping matrices. So, we define the operator $\mathfrak{Q} : \mathbb{R}^{(2k+1)\times(2k+1)} \mapsto \mathbb{R}^{(2m+1)\times(2m+1)}$, where $k \geq m$, as

$$\big(\mathfrak{Q}(\mathrm{P})\big)_{i,j} = \begin{cases} \mathrm{P}_{i,j} & \text{if } |i| < m \text{ and } |j| < m, \\ \sum_{u\geq m} \mathrm{P}_{\mathrm{sgn}(i)u,j} & \text{if } |i| = m \text{ and } |j| < m, \\ \sum_{v\geq m} \mathrm{P}_{i,\mathrm{sgn}(j)v} & \text{if } |i| < m \text{ and } |j| = m, \\ \sum_{u\geq m, v\geq m} \mathrm{P}_{\mathrm{sgn}(i)u,\mathrm{sgn}(j)v} & \text{if } |i| = m \text{ and } |j| = m, \end{cases}$$

where $\mathrm{sgn}(i)$ corresponds to the sign of i. We can now write the recursion

$$\mathrm{P}^{(t_1,t_2)} = \sum_d \lambda_d \mathfrak{Q}\big(\mathrm{C}_{\mathsf{BMS}} \odot (\mathrm{Q}^{(t_1,t_2)})^{\odot(d\text{-}1)}\big)$$
$$\mathrm{Q}^{(t_1,t_2)} = \sum_d \rho_d (\mathrm{P}^{(t_1\text{-}1,t_2\text{-}1)})^{\boxdot(d\text{-}1)},$$

with the initial condition $\mathrm{Q}^{(t_1\text{-}t_2,0)}_{i,j} = (\mathsf{b}^{(t_1\text{-}t_2)})_i \mathbb{1}_{\{j=0\}}$. Finally we have

$$\mathbb{P}\{\nu^{(\ell)} = r, \nu^{(\ell-k)} = s\} = \mathrm{P}^{(\ell,\ell-k)}_{r,s} \tag{4.1}$$

Bibliography

[1] M. E. Fisher, "Proc. of the Enrico Fermi school, Varenna, Italy, 1970, course n. 51," in *Critical Phenomena*, International School of Physics Enrico Fermi, Course LI, edited by M. S. Green. (Academic, New York, 1971), 1971.

[2] V. Privman, "Finite-size scaling theory," in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, V. Privman, Ed. World Scientific, 1990, pp. 1–98.

[3] A. Montanari, "Finite-size scaling of good codes," in *Proc. of the Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, USA, Oct. 2001.

[4] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," in *Proc. of the Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, USA, Oct. 2003.

[5] ——, "Finite-length scaling for iteratively decoded LDPC ensembles," accepted IEEE Trans. on Information Theory.

[6] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. A. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proc. of the 29th annual ACM Symposium on Theory of Computing*, 1997, pp. 150–159.

[7] A. Amraoui, A. Montanari, and R. Urbanke, "Analytic determination of scaling parameters," in *Proc. of the IEEE Int. Symposium on Inform. Theory*, Seattle, WA, July 2006, conference.

[8] ——, "How to find good finite-length codes: From art towards science," *Euro. Trans Telecomms.*, vol. 18, no. 5, pp. 491–508, Apr. 2007.

[9] J. Ezri, A. Montanari, and R. Urbanke, "A generalization of the finite-length scaling approach beyond the BEC," in *Proc. of the IEEE Int. Symposium on Inform. Theory*, Nice, France, June 2007, conference.

[10] R. G. Gallager, “Low-density parity-check codes,” *IRE Transactions on Inform. Theory*, vol. 8, pp. 21–28, jan 1962.

[11] L. Bazzi, T. Richardson, and R. Urbanke, “Exact thresholds and optimal codes for the binary-symmetric channel and Gallager’s decoding algorithm A,” *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2010–2021, Sept. 2004.

[12] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge University Press, 2008.

[13] J. Ezri, A. Montanari, S. Oh, and R. Urbanke, “Computing threshold shift for general channels,” in *Proc. of the IEEE Int. Symposium on Inform. Theory*, Toronto, Canada, July 2008, conference.

[14] A. Dembo and A. Montanari, “Finite size scaling for the core of large random hypergraphs,” 2007, e-print: math.PR/0702007.

Index