Broadcasting on Trees Near Criticality: Perturbation Theory

Qian Yu
University of Southern California

Yury Polyanskiy
Massachusetts Institute of Technology

***Abstract*—Consider a setting where a single bit is broadcast down the d-ary tree, where each edge acts as a binary symmetric channel with a crossover probability δ. The goal is to reconstruct the root bit given the values of all bits at a large distance h from the root. It is known the reconstruction is impossible iff $(1-2\delta)^2 d \leq 1$. In this paper, we show that in the regime where the latter product converges to 1 from the above, the distribution of the log-likelihood ratio (LLR) of the root bit given the far-away boundary (normalized by the square root of deviation of δ from criticality) converges to an explicit Gaussian distribution. This strengthens a similar result of Jain-Koehler-Liu-Mossel (COLT'2019) and enables us to resolve conjectures stated in Gu-Roozbehani-Polyanskiy (ISIT'2020) for the scaling of the probability of error and mutual information near criticality. Our results also provide a rationale for the ubiquitous $\mathcal{N}(\mu, 2\mu)$ approximation of the LLR distribution in the EXIT-chart heuristics.**

I. Introduction

We consider an infinite perfect d-ary tree with a root called vertex 0. Each vertex v is associated with a binary variable X_v. X_0 is Bernoulli($\frac{1}{2}$). For any vertex $v \neq 0$, let u be its parent, X_v equals X_u with a probability of $1-\delta$ and $1-X_u$ otherwise, conditioned on the collection of X'_v for all vertices $v' \neq v$ with a level less or equal to than the level of v, for some parameter[1] $\delta \in (0, \frac{1}{2}]$.

Let L_h denotes the collection of vertices with a level of h, let X_{L_h} denotes the collection of X_v for any $v \in L_h$. We are interested in the following two lists of quantities: $P(\delta, h)$ defined as the minimum probability of error for estimating X_0 using X_{L_h}; and $I(\delta, h) \triangleq I(X_0; X_{L_h})$. We aim to solve for $P(\delta) \triangleq \lim_{h\to+\infty} P(\delta, h)$ and $I(\delta) \triangleq \lim_{h\to+\infty} I(\delta, h)$.

It is known from [1]–[3] that there is a reconstruction threshold $\delta_{\mathrm{c}} = \frac{1-\frac{1}{\sqrt{d}}}{2}$ such that $P(\delta) = \frac{1}{2}$ (or equivalently, $I(\delta) = 0$) if and only if $\delta \geq \delta_{\mathrm{c}}$. However, the values of $I(\delta)$ and $P(\delta)$ are not known for $\delta < \delta_{\mathrm{c}}$. Some conjectures about the limiting behavior of these quantities as $\delta \to \delta_{\mathrm{c}}$ were stated in [4] on their decay rates and the associated multiplicative factors. In this work, we resolve those conjectures by characterizing the belief propagation (BP) fixed point distribution in this limit.

We note that a generalization of the basic setting was considered in [5], referred to as "robust reconstuction", where instead of inferring X_0 from X_{L_h}, a noisy version of the latter is observed with entries corrupted by independent and identical discrete channels. All results presented in this paper directly apply to the robust reconstruction as long as the observation channels are symmetric (or BMS [6, Chapter 4]). While we focus on a symmetric setting, broadcasting through asymmetric channels has also been studied. E.g., the first tight result is provided in [7], matching the Kesten-Stigum lower bound [2] on reconstruction threshold.

[1]For brevity, we ignore the trivial case ($\delta = 0$ or $d = 1$). Results for $\delta > \frac{1}{2}$ can be obtained by symmetry.

For convenience, we define $\tau = \delta_{\mathrm{c}} - \delta$. Note that $\tau \to 0^+$ and $\delta \to \delta_{\mathrm{c}}^-$ are equivalent for any fixed d.

A. BP recursion

To find the values of $P(\delta)$ and $I(\delta)$, we need to introduce distributional BP recursion equations. To that end, define log-likelihood ratio (LLR) distribution conditioned on $X_0 = 0$ as

$$\mathbb{P}[R_{(h)} = r] = \mathbb{P}\left[\ln \frac{P(X_{L_h}|X_0=0)}{P(X_{L_h}|X_0=1)} = r | X_0 = 0\right].$$

For $h = 0$ we set $R_{(0)} = +\infty$ w.p. 1. It is easy to check that the law μ of $R_{(h)}$ for any h satisfies the following symmetry condition:

$$d\mu(r) = e^r d\mu(-r)\,, \tag{1}$$

which for a discrete distribution is equivalent to $\mathbb{P}[R_{(h)} = r] = e^r \mathbb{P}[R_{(h)} = -r]$.

The distribution of $R_{(h)}$ can be determined recursively, as follows. Let $\tilde{R}_u$ be iid copies of $R_{(h)}$ and let $X_u \overset{iid}{\sim} (-1)^{\mathrm{Ber}(\delta)}$ (all jointly indepenent). Then

$$R_{(h+1)} \overset{(d)}{=} \sum_{u=1}^{d} X_u F_\delta(\tilde{R}_u)\,,$$

where

$$F_\delta(x) \triangleq \ln \frac{(1-\delta)e^x + \delta}{\delta e^x + 1 - \delta} = 2\tanh^{-1}((1-2\delta)\tanh\frac{x}{2})\,.$$

(The same recursion works for the robust reconstruction problem, except that $R_{(0)}$ is taken to be any general symmetric distribution, cf. (1).)

Knowing $R_{(h)}$, one can express quantities of interest as follows.

$$P(\delta, h) = \mathbb{P}[R_{(h)} < 0] + \frac{1}{2}\mathbb{P}[R_{(h)} = 0] \tag{2}$$

$$I(\delta, h) = \ln 2 - \mathbb{E}[\ln(1 + e^{-R_{(h)}})]. \tag{3}$$

Hence, to solve for $P(\delta)$ and $I(\delta)$, it suffices to characterize the distribution of $R_{(h)}$.

As $h \to \infty$ it is known that the distributions of $R_{(h)}$ converge to a distribution with the following general properties.

Definition 1 (Fixed point of BP)**.** A distribution μ is called a *BP fixed point* if given $\tilde{R}_u \overset{iid}{\sim} \mu$ and $X_u \overset{iid}{\sim} (-1)^{(\mathrm{Ber}(\delta)}$ (jointly independetly of each other) we have that

$$R \triangleq \sum_{u=1}^{d} X_u F_\delta(\tilde{R}_u)$$

also has law μ. Furthermore, we call a fixed point symmetric if μ satisfies (1) and non-trivial if $\mu[\{0\}] < 1$.

The following is well known (e.g. [8, Lemma 29]):

Proposition 1. *For each τ the distributions of $R_{(h)}$ converge to a symmetric fixed point distribution μ_τ^*, which is non-trivial iff $\tau > 0$. The same statement holds for the robust reconstruction problem with symmetric noise channels (with positive capacities) at the leaves.*

II. Main Results

Our goal is to provide new statements about the mysterious measures μ_τ^* (cf. Proposition 1) in the limit of $\tau \to 0^+$. It is widely believed that for each $\tau > 0$ there is a unique non-trivial symmetric fixed point μ_τ, however, at present this is only proved for $d(1-2\delta)^2 > 3.531$, cf. [9]. We prove, however, that unconditionally, any sequence of fixed-points μ_τ must be asymptotically Gaussian. We remind the reader of the definition of a normal family, see Chapter 8.4.2 in [10].

Theorem 1. *For any fixed d and for each τ, let μ_τ be any non-zero symmetric solution to the fixed point equation. Consider $R_\tau \sim \mu_\tau$ and let $A_\tau = R_\tau/\sqrt{\tau}$. The set of holomorphic functions $z \mapsto \mathbb{E}[e^{zA_\tau}]$ indexed by $\tau \in (0, \tau_0)$ for any $\tau_0 < \delta_c$ is uniformly bounded on any strip $\{z : |\Re(z)| < h\}$. In particular, this set forms a normal family on all of $\mathbb{C}$. Furthermore, we have*

$$\mathbb{E}[e^{zA_\tau}] \to e^{\sigma^2 z^2/2} \qquad (\text{as } \tau \to 0^+) \tag{4}$$

uniformly over any compact subset of $\mathbb{C}$, where $\sigma^2 = \lim_{\tau\to 0^+} \frac{\mathbb{E}[R_\tau^2]}{\tau} = \frac{16d\sqrt{d}}{d-1}$.

Our result imply that random variables $\frac{R_\tau}{\sqrt{\tau}}$ converge to Gaussian in distribution, and in terms of moments of all orders. We remark that previously asymptotic normality was shown in Corollary 4 of [8], albeit under a weaker mode of convergence (Wasserstein distance) and only for the special sequence μ_τ^*. Instead, our result applies to any BP-fixed point, and in particular, establishes asymptotic normality in the problem of robust reconstruction [5].

Corollary 1 (First-Order Approximations)**.** *The mutual information and probability of error are characterized by*

$$I(\delta) = \frac{2d\sqrt{d}}{d-1}\tau + o(\tau), \tag{5}$$

$$P(\delta) = \frac{1}{2} - \sqrt{\frac{2d\sqrt{d}}{\pi(d-1)}}\sqrt{\tau} + o(\sqrt{\tau}). \tag{6}$$

We compare the obtained approximations with their numerical values in Fig. 1. The first-order approximations approach the numerical values as $\tau \to 0^+$. To further reduce the gaps, we present in [11] that the distribution of R can be approximated using a sequence of density functions, which proves that $I(\delta)$ and $P(\delta)$, as well as expectations of general functions of R_τ, possess some asymptotic expansions. We also provide systematic approaches for computing them.

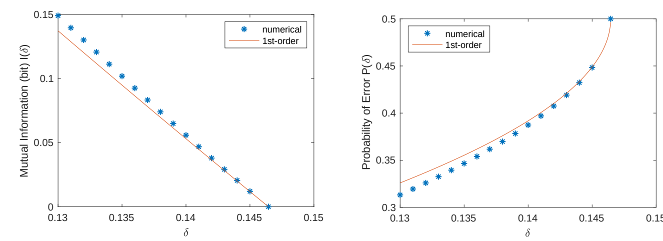

Fig. 1: Comparison between the first-order approximations and the numerical values, for $I(\delta)$ and $P(\delta)$.

Remark 1. Theorem 1 states that $\mathcal{N}(0, \tau\sigma^2)$ approximates the distribution of R_τ. However, this distribution does not satisfy (1). Its symmetrized version would be $\mathcal{N}(\frac{\tau}{2}\sigma^2, \tau\sigma^2)$, as often used in the EXIT-chart method [6, Section 4.10]. Table I shows that moments of the latter indeed better approximate moments of $R_\tau = \sqrt{\tau}A_\tau$. In fact, in [11] we generalize this result by showing that, while $\mathbb{E}[R_\tau^{2k}] \asymp \tau^k$, the $2k$-th cumulant $\kappa_{2k}(R_\tau) \asymp \tau^{2k-1}$ decaying at twice the speed.

Distributions	$2k$-th moment	$(2k-1)$-th moment
$\mathcal{N}(0, \tau\sigma^2)$	$\tau^k\sigma^{2k}(2k-1)!!$	0
$\mathcal{N}(\frac{\tau}{2}\sigma^2, \tau\sigma^2)$	$\tau^k\sigma^{2k}(2k-1)!! + o(\tau^{k+1})$	$\frac{\tau^k\sigma^{2k}(2k-1)!!}{2} + o(\tau^{k+1})$
R_τ	$\tau^k\sigma^{2k}(2k-1)!! + o(\tau^{k+1})$	$\frac{\tau^k\sigma^{2k}(2k-1)!!}{2} + o(\tau^{k+1})$

TABLE I: A comparison of moments of different distributions.

Remark 2. The Gaussian convergence is intuitive from the fixed point equation. Note that $F_\delta(R)$ is approximately linear when R is small. Particularly, $F_\delta(R) = R/\sqrt{d} + O(\tau R + R^3)$. Hence, by approximating it using the linear term $R/\sqrt{d}$, and assuming $R \overset{(d)}{\approx} -R$ from the symmetry condition. The fixed point equation gives the following for any k,

$$R \overset{(d)}{\approx} \frac{\sum_{u=1}^{d} \tilde{R}_u}{\sqrt{d}} \overset{(d)}{\approx} \frac{\sum_{u=1}^{d^k} \tilde{R}_u}{\sqrt{d^k}},$$

leading to normal distributions by central limit theorem.

III. Proof of Gaussian Limit

We prove Theorem 1 and Corollary 1 in this section. We denote by R_τ a random variable with distribution μ_τ, which is a non-trivial symmetric BP fixed point. The notation $o(1)$, $O(1)$, $O(\tau^k)$ etc all refer to the regime of $\tau \to 0^+$.

A. Convergence of moments

To demonstrate the main proof ideas, we first compute the 0-th order expansion and also show applicability of Taylor series expansion. I.e. we show that as $\tau \to 0$, $\mathbb{E}[R_\tau^2] = o(1)$ and $\mathbb{E}[R_\tau^{2k}] = O(\mathbb{E}[R_\tau^2]^k)$, making Taylor series expansion convergent (this is what is known as "perturbation theory" in physics). Note that F_δ is bounded, any solution to the fixed point equation is also bounded. More precisely, we have the following proposition.

Proposition 2 (Boundedness of R)**.** *The distribution of R_τ satisfies*

$$\mathbb{P}[|R_\tau| \le R_{\mathbb{1}}] = 1 \tag{7}$$

where $R_{\mathbb{1}} \triangleq d \ln \frac{1-\delta}{\delta} = O(1)$.

Hence, $\mathbb{E}[f(R_\tau)]$ is well-defined for any bounded Borel f. We also have the following rules by symmetry condition.

Proposition 3 (Symmetrization and comparison rules)**.** *For any bounded Borel $f : [-R_{\mathbb{1}}, R_{\mathbb{1}}] \to \mathbb{R}$ we have*

$$\mathbb{E}[f(R_\tau)] = \mathbb{E}\left[\frac{e^{\frac{1}{2}R_\tau} f(R_\tau) + e^{-\frac{1}{2}R_\tau} f(-R_\tau)}{e^{\frac{1}{2}R_\tau} + e^{-\frac{1}{2}R_\tau}}\right]. \tag{8}$$

In particular, for any odd function f we have $\mathbb{E}[f(R_\tau)] = \mathbb{E}[f(R_\tau)\tanh\frac{R_\tau}{2}]$. Consequently, if f and g are two odd functions with $f \ge g$ on $[0, R_{\mathbb{1}}]$ then

$$\mathbb{E}[f(R_\tau)] \ge \mathbb{E}[g(R_\tau)]. \tag{9}$$

Proof. Let $\tilde{f}(R_\tau) \triangleq \frac{e^{-\frac{1}{2}R_\tau} f(R_\tau)}{e^{\frac{1}{2}R_\tau} + e^{-\frac{1}{2}R_\tau}}$. Equation (8) is implied by $\mathbb{E}[\tilde{f}(R_\tau)] = \mathbb{E}[\tilde{f}(-R_\tau)e^{-R_\tau}]$, which is due to the symmetry condition of μ_τ. □

We start by showing a simple result characterizing the reconstruction threshold δ_c.

Lemma 1. *For $\tau < 0$ we have $\mathbb{E}[R_\tau^2] = 0$, for $\tau > 0$ we have $\mathbb{E}[R_\tau^2] = o(1)$ as $\tau \to 0+$.*

Proof. According to the fixed point equation, by taking the expectation of R,

$$\mathbb{E}[R_\tau] = d(1-2\delta)\mathbb{E}\left[F_\delta(R_\tau)\right]. \tag{10}$$

Note that F_δ is an odd function, and it is easy to show that due to (7):

$$F_\delta(R_\tau) \le F'_\delta(0)R_\tau - cR_\tau^3 = (1-2\delta)R_\tau - cR_\tau^3,$$

for $c = \frac{1-2\delta}{6}\min\{\delta(1-\delta), \frac{1}{R_{\mathbb{1}}^2}\}$. Together with (10) we get

$$\mathbb{E}[R_\tau] \le d(1-2\delta)^2\mathbb{E}[R_\tau] - c\mathbb{E}[R_\tau^3]$$

By symmetrization, we have $\mathbb{E}[R_\tau^3] = \mathbb{E}[R_\tau^3 \tanh(\frac{1}{2}R_\tau)] \ge \frac{1-2\delta}{R_{\mathbb{1}}}\mathbb{E}[R_\tau^4] \ge \frac{1-2\delta}{R_{\mathbb{1}}}\mathbb{E}[R_\tau^2]^2$. Hence,

$$\mathbb{E}[R_\tau] \le d(1-2\delta)^2\mathbb{E}[R_\tau] - c(1-2\delta)\frac{\mathbb{E}[R_\tau^2]^2}{R_{\mathbb{1}}},$$

which is, equivalently,

$$((1-2\delta)^2 - \frac{1}{d})\mathbb{E}[R_\tau] \ge c(1-2\delta)\frac{\mathbb{E}[R_\tau^2]^2}{R_{\mathbb{1}}}.$$

Because $\mathbb{E}[R_\tau] \ge 0$ (e.g. from (9)), if $(1-2\delta)^2 - \frac{1}{d} \le 0$, i.e., $\delta \ge \delta_c$, we have $\mathbb{E}[R_\tau^2] = 0$. This corresponds to the case that $I(\delta) = 0$ and $P(\delta) = \frac{1}{2}$, recovering the well known reconstructability result. When $\tau > 0$ we have $(1-2\delta)^2 - \frac{1}{d} \asymp \tau$, yielding $\mathbb{E}[R_\tau^2] = O(\sqrt{\tau}) = o(1)$. □

For higher moments, the results are obtained similarly by expanding the expectation of R_τ^{2k} using the fixed point equation.

Lemma 2. *For $\tau \to 0^+$ and any integer $k > 0$, we have*

$$\mathbb{E}[R_\tau^{2k}] = O(\mathbb{E}[R_\tau^2]^k). \tag{11}$$

Furthermore, we have

$$\mathbb{E}[R_\tau^{2k}] = (2k-1)!!\mathbb{E}[R_\tau^2]^k(1+o(1)). \tag{12}$$

Proof. We prove Lemma 2 by induction. The $k = 1$ case is trivial, so we focus on $k > 1$.

We start by proving equation (11). For any $k > 1$, we evaluate $\mathbb{E}[R_\tau^{2k}]$ using the fixed point equation.

$$\begin{aligned}\mathbb{E}[R_\tau^{2k}] =& \mathbb{E}\left[\left(\sum_{u=1}^d X_u F_\delta(\tilde{R}_u)\right)^{2k}\right] \\ =& \sum_{m_1+m_2+\ldots+m_d=2k}\binom{2k}{m_1,\ldots,m_d} \\ & \prod_{u=1}^d \mathbb{E}\left[\left(X_u F_\delta(\tilde{R}_u)\right)^{m_u}\right]. \end{aligned} \tag{13}$$

From $F_\delta(x) \le (1-2\delta)x$ for $x > 0$ we obtain for any even m:

$$\mathbb{E}[(X_u F_\delta(\tilde{R}_u))^m] \le (1-2\delta)^m \mathbb{E}[R_\tau^m]. \tag{14}$$

Consequently, from the induction hypothesis, every term in (13) with m_u–even and $< 2k$ is of order $\mathbb{E}[R_\tau^2]^k$. We next show that all other terms with all $m_u < 2k$ are $o(\mathbb{E}[R_\tau^2]^k)$ or $o(\mathbb{E}[R_\tau^{2k}])$. First, we notice the estimate for any odd m via (8):

$$\begin{aligned}\mathbb{E}[(X_u F_\delta(\tilde{R}_u))^m] &= (1-2\delta)\mathbb{E}\left[(F_\delta(R_\tau))^m \tanh\frac{R_\tau}{2}\right] \\ &= O(\mathbb{E}[R_\tau^{m+1}])\end{aligned} \tag{15}$$

Next, consider terms where $m_u = 2k-1$ (and thus some other $m_{u'} = 1$). For such terms, we have $\prod_u \mathbb{E}[R_\tau^{m_u}] = O(\mathbb{E}[R_\tau^{2k}]\mathbb{E}[R_\tau^2])$ which is $o(\mathbb{E}[R_\tau^{2k}])$ via Lemma 1. Next, consider terms with all $m_u < 2k-1$. Then, for every odd m_u from (15) and induction hypothesis we get $\mathbb{E}[R_\tau^{m_u+1}] \asymp \mathbb{E}[R_\tau^2]^{\frac{m_u+1}{2}}$. Thus, taking the product of such terms we get overall order of $\mathbb{E}[R_\tau^2]$ to be strictly greater than k (recall $\sum_u m_u = 2k$). Putting everything together, we obtained:

$$\begin{aligned}\mathbb{E}[R_\tau^{2k}] \le& o(\mathbb{E}[R_\tau^{2k}]) + o(\mathbb{E}[R_\tau^2]^k) + (1-2\delta)^{2k} \\ &\cdot \sum_{\substack{m_1+m_2+\ldots+m_d=2k \\ 2|m_u,\ \forall u\in[d]}}\binom{2k}{m_1,\ldots,m_d}\prod_{u=1}^d \mathbb{E}[R_\tau^{m_u}].\end{aligned} \tag{16}$$

We apply the induction assumption.

$$\mathbb{E}[R_\tau^{2k}] \le (1-2\delta)^{2k} d\mathbb{E}[R_\tau^{2k}] + O(\mathbb{E}[R_\tau^2]^k) + o(\mathbb{E}[R_\tau^{2k}]).$$

Recall that $\mathbb{E}[R_\tau^2] = o(1)$.

$$\begin{aligned}\mathbb{E}[R_\tau^{2k}] \leq & \frac{1}{1-(1-2\delta)^{2k}d - O(\mathbb{E}[R_\tau^2])} O(\mathbb{E}[R_\tau^2]^k) \\ = & O(\mathbb{E}[R_\tau^2]^k). \end{aligned} \tag{17}$$

To prove (12), we can rederive (16) as equality by improving (14). Indeed, it is easy to show that due to (7) we have $F_\delta(R_\tau) = (1-2\delta)R_\tau + c(R_\tau)R_\tau^3$ where $|c(R_\tau)|$ is bounded. By raising this identity to the even power m and repeatedly applying (11) we obtain

$$\begin{aligned}\mathbb{E}[(X_u F_\delta(\tilde{R}_u))^m] = \mathbb{E}[F_\delta(\tilde{R}_u)^m] &= (1-2\delta)^m \mathbb{E}[R_\tau^m](1+o(1)) \\ &= \left(\frac{1}{\sqrt{d}}\right)^m \mathbb{E}[R_\tau^m](1+o(1)) \end{aligned} \tag{18}$$

Consequently, we obtain the following expression for $\mathbb{E}[R_\tau^{2k}]$:

$$\sum_{\substack{m_1+\ldots+m_d=2k \\ 2|m_u,\ \forall u\in[d]}} \binom{2k}{m_1,\ldots,m_d} \prod_{u=1}^d \mathbb{E}\left[\left(\frac{R_\tau}{\sqrt{d}}\right)^{m_u}\right](1+o(1)).$$

Now consider a Gaussian variable $G \sim \mathcal{N}(0, \mathbb{E}[R_\tau^2])$. Then by infinite divisibility of the Gaussian we have

$$\mathbb{E}\left[G^{2k}\right] = \mathbb{E}\left[\left(\frac{G_1}{\sqrt{d}} + \cdots + \frac{G_d}{\sqrt{d}}\right)^{2k}\right] \tag{19}$$

Since odd moments $\mathbb{E}[G^m] = 0$ after expanding the power, we obtain exactly the same expansion as the dominant terms for $\mathbb{E}[R_\tau^{2k}]$. Since by the induction assumption we already know $\mathbb{E}[R_\tau^{2k_1}] = \mathbb{E}[G^{2k_1}](1+o(1))$ for all $k_1 < k$, we conclude that also

$$\mathbb{E}[R_\tau^{2k}] = \mathbb{E}[G^{2k}](1+o(1)) = (2k-1)!!\mathbb{E}[R_\tau^2]^k(1+o(1)).$$

□

B. First-Order Approximation

Lemma 3. *For $\tau < 0$ we have $\mathbb{E}[R_\tau^2] = 0$. For $\tau \to 0+$ we have $\mathbb{E}[R_\tau^2] = O(\tau)$, and more precisely,*

$$\mathbb{E}[R_\tau^2] = \frac{16d\sqrt{d}\tau}{d-1} + o(\tau). \tag{20}$$

Remark 3. The fact that $\mathbb{E}[R_\tau^2] \asymp \tau$ can be guessed as follows. Note that there is a $r^*(\tau) > 0$ such that $|F_\delta(R)| > |R/\sqrt{d}|$ $\iff$ $|R| < r^*(\tau)$. Hence, the nonlinearity of $F_\delta(R)$ favors further reducing $\mathbb{E}[R_\tau^2]$ when $|R_\tau|$ is large, and vice versa. As a consequence, the scale of R_τ tends to be stabilized at $r^*(\tau)$, which is at the level of $\sqrt{\tau}$.

Proof. We use the fixed point equation to evaluate the second moment.

$$\mathbb{E}[R_\tau^2] = d\mathbb{E}\left[F_\delta(R_\tau)^2\right] + d(d-1)\left((1-2\delta)\mathbb{E}\left[F_\delta(R_\tau)\right]\right)^2.$$

By expanding and approximating F_δ with Lemma 2 and symmetrization, we can express both $\mathbb{E}\left[F_\delta(R_\tau)\right]^2$ and $\mathbb{E}\left[F_\delta(R_\tau)^2\right]$ using $\mathbb{E}[R_\tau^2]$, with an error term up to $o(\mathbb{E}[R_\tau^2]^2)$.

$$\begin{aligned}\mathbb{E}\left[F_\delta(R_\tau)\right]^2 =& \frac{(1-2\delta)^2}{4}\mathbb{E}[R_\tau^2]^2 + o(\mathbb{E}[R_\tau^2]^2) \\ \mathbb{E}\left[F_\delta(R_\tau)^2\right] =& (1-2\delta)^2(\mathbb{E}[R_\tau^2] - 2\delta(1-\delta)\mathbb{E}[R_\tau^2]^2) \\ &+ o(\mathbb{E}[R_\tau^2]^2)\end{aligned}$$

Then by $d(1-2\delta)^2 = 1 + o(1)$, we have

$$(d(1-2\delta)^2 - 1)\mathbb{E}[R_\tau^2] = \frac{(d-1)}{4d}\mathbb{E}[R_\tau^2]^2(1+o(1)). \tag{21}$$

Recall the non-zero requirement. We have $\mathbb{E}[R_\tau^2] > 0$. Then the above quadratic equation implies that

$$\begin{aligned}\mathbb{E}[R_\tau^2] &= \frac{4d}{d-1}(d(1-2\delta)^2 - 1)(1+o(1)) \\ &= \frac{16d\sqrt{d}\tau}{d-1} + O(\tau).\end{aligned}$$

□

C. Subgaussianity of $\frac{R_\tau}{\sqrt{\tau}}$

Theorem 1 also relies on the following final ingredient.

Lemma 4. *In the conditions of Theorem 1, define $M_{R_\tau}(s) = \mathbb{E}[e^{sR_\tau}]$. Then $\forall \delta_0 \in (0, \delta_c)$, there are C_1, $C_2 > 0$ such that for any $z \in \mathbb{C}$ and $\tau < \delta_c - \delta_0$, we have*

$$|M_{R_\tau}(z/\sqrt{\tau})| \leq C_1 \exp(C_2 \Re\{z\}^2).$$

Proof. Recall the definition of moment-generating functions. We have $|M_{R_\tau}(z/\sqrt{\tau})| \leq M_{R_\tau}(\Re\{z\}/\sqrt{\tau})$. It suffices to prove for the case where $z \in \mathbb{R}$.

We first prove the following unconditional estimate. For every $s \in \mathbb{R}$ and integer $r \in \mathbb{N}$:

$$\ln M_{R_\tau}(s) \leq d^r \ln M_{R_\tau}((1-2\delta)^r |s|). \tag{22}$$

By symmetrization, for any odd functions f, g, s.t. $f \geq |g|$ on $\mathbb{R}_+$, we have

$$\begin{aligned}&\mathbb{E}[e^{f(R_\tau)}] - \mathbb{E}[e^{g(R_\tau)}] = \\ &\mathbb{E}\left[\frac{\cosh\left(\frac{R_\tau}{2} + f(R_\tau)\right) - \cosh\left(\frac{R_\tau}{2} + g(R_\tau)\right)}{\cosh\frac{R_\tau}{2}}\right] \geq 0.\end{aligned} \tag{23}$$

Consequently, for any $s \geq 0$, $\mathbb{E}[e^{-sR_\tau}] \leq \mathbb{E}[e^{sR_\tau}]$, and $\mathbb{E}[e^{-sF_\delta(R_\tau)}] \leq \mathbb{E}[e^{sF_\delta(R_\tau)}] \leq \mathbb{E}[e^{s(1-2\delta)R_\tau}]$. In all, this implies

$$\mathbb{E}[e^{sR_\tau}] \leq \mathbb{E}[e^{|s|F_\delta(R_\tau)}]^d \leq \mathbb{E}[e^{|s|(1-2\delta)R_\tau}]^d.$$

By applying the above inequality recursively, we have for any integer $r \geq 0$,

$$\mathbb{E}[e^{sR_\tau}] \leq \mathbb{E}[|e^{|s|(1-2\delta)^r R_\tau}|]^{d^r}, \tag{24}$$

which is essentially inequality (22).

Second, consider any δ_0 bounded away from zero. Due to symmetrization and boundedness of R, we can uniformly upper bound $\ln M_{R_\tau}(s)$ for $|s| \le 1$ by $O(\mathbb{E}[R_\tau^2])$. In particular, we can find $c_1 > 0$ such that for any $\tau < \delta_c - \delta_0$ and $|s| \le 1$,

$$\ln M_{R_\tau}(s) \le \ln(1 + c_1 \mathbb{E}[R_\tau^2]) \le c_1 \mathbb{E}[R_\tau^2], \tag{25}$$

Then for general $|s| \in [1, \frac{1}{\tau}]$, the needed result for this regime can be obtained by applying inequality (22). Note that we can find integer r such that $|s|(1-2\delta)^r \le 1$ with $r \le 1 + \frac{\ln |s|}{|\ln(1-2\delta)|}$. By applying inequality (22), we have

$$\ln M_{R_\tau}(s) \le d^{\frac{\ln |s|}{|\ln(1-2\delta)|}+1} c_1 \mathbb{E}[R_\tau^2] = |s|^{\frac{\ln d}{|\ln(1-2\delta)|}} c_1 d \mathbb{E}[R_\tau^2].$$

Note that $\frac{\ln d}{|\ln(1-2\delta)|} \le 2 + \frac{2\sqrt{d}}{|\ln(1-2\delta_0)|}\tau$. For any $|s| \in [1, 1/\tau]$,

$$\begin{aligned}\ln M_{R_\tau}(s) &\le e^{\frac{2\sqrt{d}}{|\ln(1-2\delta_0)|}\tau|\ln \tau|} c_1 d \mathbb{E}[R_\tau^2] s^2 \\ &\le e^{\frac{2\sqrt{d}}{e|\ln(1-2\delta_0)|}} c_1 d \mathbb{E}[R_\tau^2] s^2 \end{aligned} \tag{26}$$

Combining inequalities (25) and (26), there is $C > 0$ such that for any $|s| \le 1/\tau$ and $\tau < \delta_c - \delta_0$,

$$\ln M_{R_\tau}(s) \le C\mathbb{E}[R_\tau^2](1 + s^2). \tag{27}$$

Using the fact that $\mathbb{E}[R_\tau^2] = O(\tau)$ and R_τ is bounded, we can find c_2, $c_3 > 0$ such that for any $\tau < \delta_c - \delta_0$, $\mathbb{E}[R_\tau^2] \le c_2\tau$ and $R_\mathbb{1} \le c_3$. Hence,

$$\ln M_{R_\tau}(s) \le C(c_3 + c_2 s^2 \tau). \tag{28}$$

Finally, using the boundedness of R, for any $|s| \ge 1/\tau$ and $\tau < \delta_c - \delta_0$, we can find $R_\mathbb{1} < c$. Then,

$$\ln M_{R_\tau}(s) \le \ln e^{c|s|} \le cs^2\tau \tag{29}$$

The proof is concluded by combining inequalities (28) and (29), which collectively cover the entire real line.

□

D. Proofs of the Theorem and Corollary

Proof of Theorem 1. From Lemma 4, the family of functions $z \mapsto \mathbb{E}[e^{zA_\tau}]$ for any $\tau < \delta_c - \delta_0$ is uniformly bounded on any compact subset of $\mathbb{C}$. Hence, they form a normal family on $\mathbb{C}$.

According to Lemma 2 and symmetry condition, all derivatives at zero of the moment-generating function $f_\tau(z) = \mathbb{E}[e^{zA_\tau}]$ converge to those of $f(z) = e^{\frac{\sigma^2 z^2}{2}}$, corresponding to $\mathcal{N}(0, \sigma^2)$. This implies that this family may have at most one limit point: $f(z)$. (Indeed, every limit point must be a holomorphic function itself by Cauchy's formula, with its Maclaurin series coinciding with that of f.) Suppose $f_\tau \not\to f$. Then on some compact $K \subset \mathbb{C}$ we have a sequence $\sup_{z\in K} |f_{\tau_n}(z) - f(z)| \ge \epsilon_0 > 0$ for all n. But by normality of the family, the subsequence f_{τ_n} must have a limit point and by the argument above it must be f, a contradiction. □

We continue to characterize $I(\delta)$ and $P(\delta)$ for Corollary 1 using the proved Gaussian Convergence.

Proof of Corollary. By symmetrization rule and (3), we have

$$I(\delta) = \mathbb{E}\left[\frac{R_\tau}{2}\tanh\frac{R_\tau}{2} - \ln\cosh\frac{R_\tau}{2}\right], \tag{30}$$

which is expectation of an even C^∞ function of R_τ. Using Lemma 2, it can be expanded and approximated using even moments of R_τ.

$$I(\delta) = \frac{1}{8}\mathbb{E}[R_\tau^2] + O(\mathbb{E}[R_\tau^4]) = \frac{1}{8}\mathbb{E}[R_\tau^2] + O(\mathbb{E}[R_\tau^2]^2). \tag{31}$$

Then equation (5) follow from Lemma 3.

For $P(\delta)$, first we need to show that it can be characterized as the expectation of a function of R. Because $P(\delta)$ is defined based on the expectation of a step function from a recursion process. Unlike $I(\delta)$, its connection to the limiting distribution is not directly implied by weak convergence. This can be handled by applying symmetrization rule to the recursion, and we get

$$P(\delta) = \mathbb{E}\left[\frac{e^{-\frac{1}{2}|R_\tau|}}{e^{\frac{1}{2}R_\tau} + e^{-\frac{1}{2}R_\tau}}\right]. \tag{32}$$

Next, by expanding exponents in Taylor series (and using the uniform bound $|R_\tau|^3 \le CR_\tau^2$ for some constant $C > 0$) we obtain

$$P(\delta) = \frac{1}{2} - \frac{\mathbb{E}[|R_\tau|]}{4} + O(\mathbb{E}[R_\tau^2]). \tag{33}$$

Recalling that $R_\tau = \sqrt{\tau}A_\tau$, it suffices to find the 1st-order approximation for $\mathbb{E}[|A_\tau|]$. From Theorem 1 we know that the family $\{A_\tau, \tau \in (0, \tau_0)\}$ for any $\tau_0 < \delta_c$ is uniformly integrable. By Skorokhod representation we may assume $A_\tau \overset{(a.s.)}{\to} A_0 = \mathcal{N}(0, \sigma^2)$. Then from uniform integrability we have $A_\tau \overset{L_1}{\to} A_0$ and, in particular, $\mathbb{E}[|A_\tau|] \to \mathbb{E}[|A_0|] = \sqrt{\frac{2\sigma^2}{\pi}}$. □

IV. Higher-Order Corrections

In [11], we show that the distribution of R can be approximated in a form of asymptotic expansions, up to any degree. In particular, we proved that all moments and cumulants of R can be written into the forms of $C_1\tau^k + C_2\tau^{k+1} + ...$, where any of the coefficients can be computed as a closed-form function of d. Moreover, we show that the cumulants of R decay twice faster compared to the moments, providing a stronger characterization than Gaussian convergence. They imply the uniqness of the non-zero symmetric solutions to the fixed point equation near criticality up to an error term super-polynomial in τ.

Generally, we prove the existence of a sequence of density functions, each simply given by a polynomial of R multiplied by a normal density function, which can be used to approximate the expectation of any function of R that is Lebesgue integrable on compact sets. In turn, these results imply higher-order expansions of $I(\delta)$ and $P(\delta)$, and provide better approximations near $\tau \to 0$, as illustrated in Fig 2.

References

[1] P. M. Bleher, J. Ruiz, and V. A. Zagrebnov, "On the purity of the limiting gibbs state for the ising model on the bethe lattice," *Journal of Statistical Physics*, vol. 79, no. 1, pp. 473–482, 1995. [Online]. Available: https://doi.org/10.1007/BF02179399

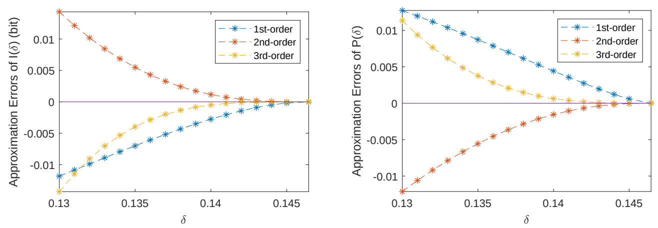

Fig. 2: Comparison of approximation errors of different orders for $d = 2$, defined as the difference between the truncated asymptotic expansions and the exact values. The data points for exact values are obtained numerically.

[2] H. Kesten and B. P. Stigum, "Additional Limit Theorems for Indecomposable Multidimensional Galton-Watson Processes," *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1463 – 1481, 1966. [Online]. Available: https://doi.org/10.1214/aoms/1177699139

[3] T. Moore and J. L. Snell, "A branching process showing a phase transition," *Journal of Applied Probability*, vol. 16, no. 2, p. 252–260, 1979.

[4] Y. Gu, H. Roozbehani, and Y. Polyanskiy, "Broadcasting on trees near criticality," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 1504–1509.

[5] S. Janson and E. Mossel, "Robust reconstruction on trees is determined by the second eigenvalue," *Ann. Probab.*, vol. 32, no. 3B, pp. 2630–2649, 07 2004. [Online]. Available: https://doi.org/10.1214/009117904000000153

[6] T. Richardson and R. Urbanke, *Modern coding theory*. Cambridge university press, 2008.

[7] C. Borgs, J. Chayes, E. Mossel, and S. Roch, "The kesten-stigum reconstruction bound is tight for roughly symmetric binary channels," in *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, 2006, pp. 518–530.

[8] V. Jain, F. Koehler, J. Liu, and E. Mossel, "Accuracy-memory tradeoffs and phase transitions in belief propagation," ser. Proceedings of Machine Learning Research, A. Beygelzimer and D. Hsu, Eds., vol. 99. Phoenix, USA: PMLR, 25–28 Jun 2019, pp. 1756–1771. [Online]. Available: http://proceedings.mlr.press/v99/jain19b.html

[9] E. Abbe, C. Elisabetta, Y. Gu, and Y. Polyanskiy, "Stochastic block model entropy and broadcasting on trees with survey," *arXiv preprint arXiv:2101.12601*, Jan. 2021.

[10] S. G. Krantz, *Handbook of complex variables*. Basel: Birkhäuser, 1999.

[11] Q. Yu and Y. Polyanskiy, "Ising model on a tree: Uniqueness of BP fixed point and analysis near criticality," *online preprint*, 2021.