Fractal Geometry of Minimal Spanning Trees

T. S. Jackson and N. Read

Part II: arXiv:0909.5343
Accepted by Phys. Rev. E
Problem Definition

- Given a graph with costs on the edges, the MST is
  - A tree (no loops)
  - Spanning: all vertices on the tree
  - Minimum: the spanning tree which minimizes total cost of edges on the tree (= energy)
Problem Definition

• Given a graph with costs on the edges, the MST is
  • A tree (no loops)
  • Spanning: all vertices on the tree
  • Minimum: the spanning tree which minimizes total cost of edges on the tree (= energy)
Problem Definition

• Given a graph with costs on the edges, the MST is
  • A tree (no loops)
  • Spanning: all vertices on the tree
  • Minimum: the spanning tree which minimizes total cost of edges on the tree (= energy)
Relevant properties

Invariance under reparameterization of costs

MST = Union of max. cost edge (barrier)-minimizing paths (geodesics)

e.g.,

Construction via greedy algorithms
(no metastability, no RSB)
Motivation

• MST paths = solutions to strong disorder transport problems
  • “Strong disorder” = extensive variance
  • E.g., resistor networks, polymers

• Nontrivial results (exponents) for a model with quenched disorder
  • Equivalently: nonlocal optimization problem in terms of local field theory

• Newman and Stein ‘94: mapped EA spin glass ground state in strong disorder limit to MSTs
Objects of study

$D$ of MST paths via \textit{path vertex} connectedness function

$D$ of MSF clusters visible in window → $d_c$, via hyperscaling breakdown
(→ SD spin glass ground state enumeration)
Connection to Percolation

Disorder: edge costs are IID quenched random variables uniformly distributed on $[0,1]$

Kruskal’s greedy algorithm for MST: add cheapest remaining edge unless a loop forms

Bond percolation: raise occupation probability, always accept cheapest remaining edge

$p_0$
Rest of Talk

• **Part I**: Mean field theory on Bethe lattice
  • MFT breaks down at $d_c = 6$
  • How? MSF cluster proliferation from multiple percolation spanning clusters (e.g., Aizenman ’97)

• **Part II**: Loop corrections for $d < d_c$
  • MSF cluster $D = d$, so look at MST paths
  • Adapt field theoretic perturbation series for percolation
  • Prove series is renormalizable to *all orders*
  • 1-loop example RG calculation
Part I: Mean field theory

No finite cycles
Bethe Lattice percolation

Coordination $z = \sigma + 1 = 3$
Bethe Lattice percolation

Coordination \( z = \sigma + 1 = 3 \)

Percolation:
\[
F(p) = 1 - p + pF^\sigma(p)
\]

\( p_c = \frac{1}{\sigma} \)

\[ F(p) \]

\[ p \]
Bethe Lattice MSTs

MST:
Bethe Lattice MSTs

MST:

- Use wired boundary conditions
Bethe Lattice MSTs

MST:
- Use wired boundary conditions
- MSF in bulk
Bethe Lattice MSTs

MST:
• Use wired boundary conditions
• MSF in bulk
• Q: Are two points connected in the bulk or via boundary?
Bethe Lattice MSTs

MST:
- Use wired boundary conditions
- MSF in bulk
- Q: Are two points connected in the bulk or via boundary?
  - Mean-field $D$ of MSF clusters
**Bethe Lattice MSTs**

**MST:**
- Use wired boundary conditions
- MSF in bulk
- Q: Are two points connected in the bulk or via boundary?
  - $= \text{Mean-field } D$ of MSF clusters
  - $= d_c$ where hyperscaling, MFT break down
MFT connectedness functions

Definitions:

\[ P_{(0;m)} = \Pr_{0}^{m} = \Pr \]
MFT connectedness functions

Definitions:

\[ P_{0;m} = \begin{array}{c}
0 \\
m
\end{array} = \begin{array}{cccc}
l_{0,\infty} & & & \\
& l_{j-1,\infty} & l_{j,\infty} & l_{m,\infty}
\end{array} \]

Boundary
MFT connectedness functions

Definitions:

\[ P_{(0;m)} = \begin{array}{l}
0 \\
\hline
m
\end{array} \]

Work with rates \( \Phi_j(p) = \frac{d}{dp} P'_{(0;j)}(p) \)

Recursion:

\[ \Phi_j(p) = F(p)^{\sigma-1} \left( p\Phi_{j-1}(p) + \int_0^p dp' \Phi_{j-1}(p') \right) \]

Either

\[ = \int_0^1 dp' \left[ F(p)^{\sigma-1} \frac{d}{dp} [p\theta(p - p')] \right] \Phi_{j-1}(p') \]

\[ \equiv K(p, p') \]
MFT connectedness functions

Connectedness functions via iterated kernel:

\[ K^m(p, p') = \int_0^1 dp_1 \cdots \int_0^1 dp_{m-1} K(p, p_1) \cdots K(p_{m-1}, p') \]

= real-space propagator

Initial condition: \( \Phi_0(p) = \frac{d}{dp} (1 - F(p)\sigma) = \Phi_0(p) \)

Then \( P_{(0; m)}(p_0) = \int_0^{p_0} dp \, dp' \, \Phi_0(p') K^m(p, p') F(p) \)

and \( C^{(2)}(m; p_0) = \sum_{m'=0}^{m} F_{m', m-m'} K^m(p, p') F(p) \)

& similarly for n-point functions

= tree-level diagrams
Kernel asymptotics

Generating function

\[ \tilde{\Phi}_w(p) = \sum_{j=0}^{\infty} w^j \Phi_j(p) \]

“discrete Fourier-Laplace transform”

Invert via Green’s function \( g_w(p,p') \)

\[ = \text{momentum space propagator} \]

\[ \int_0^1 dp_2 \left[ \delta(p_1 - p_2) - wK(p_1, p_2) \right] g_w(p_2, p_3) = \delta(p_1 - p_3) \]

\[ K^*m(p_1, p_2) = \frac{1}{2\pi i} \int dw \frac{g_w(p_1, p_2)}{w^{m+1}} \]

Extract large-\( m \) behavior from first pole encountered at \( w_c \)

\[ 1 - w_c p_2 F^{\sigma-1}(p_2) = 0 \]
MFT scaling from lattice

“1-pt.”: \[ P_{(0,m)} = \frac{P_\infty(p_0)}{(\sigma + 1)\sigma^{m-1}} \rightarrow \sum_{\text{sites}} P_{(0,m')} (p_0) = mP_\infty(p_0) \]

Compare generating function w/Fourier: \( m \sim r^2 \)

⇒ Paths on MSF look like random walks

Asymptotic regime: \( m \gg \mathcal{O}((p_0 - p_c)^{-1}) \) “correlation length”

2-pt.: \[ C^{(2)}(m; p_0) \sim \frac{1}{3}(\sigma \delta p_0)^2 m^2 \sigma^{-m} \rightarrow \sum_{\text{sites}} C^{(2)}(m'; p_0) \sim m^3 \delta p_0^2 \]

⇒ MSF clusters have fractal dimension \( D = 6 \)

⇒ \( d_c = 6 \) for MSTs

\((k+1)\)-pt.: \[ \sum_{\text{sites}} C^{(k+1)} = \overline{M(m)^k} \sim m^{3k} \Rightarrow \text{not multifractal} \]
Part II: Kruskal clusters

at $d < d_c$, $p \leq p_c$

Only finite cycles
Diagrammatics

- Percolation = $q \to 1$ Potts model; want *geometric* interpretation

- Return to low-$p$ expansion for percolation on any finite graph (*Essam* ‘72, ‘80)

- Percolation: two points connected if there exists at least one path between them
  - Expand using inclusion-exclusion on paths
  - Group all subsets of paths w/same support $\to$ sum over *subgraphs*

- Equivalent form using subsets of edges
- Graphs carry associated diagrammatic weight identical to $q \to 1$ Potts tensor contractions
Diagrammatics

- MSF problem: want $\Pr[x,y \text{ connected by path through } z]$
- Can identify MST path between two points as minimum of ordering relation
- Consequence of geodesic property $\rightarrow$ ultrametricity (transport applications)
- Weight now depends on barrier cost ordering $\pi$
- Inclusion-exclusion still applicable; must sum over orderings of each graph

$$\tilde{C}_{x,y}^z(\ell_0) = \sum_{G \in G_{x,y,z}} \Pr[G \leq \ell_0] \sum_{\pi \in S_{|E(G)|}} d_{\text{MSF}}(G|\pi) \Pr[\pi],$$
**Example calculation**

**Percolation (3-pt)**

\[
d(G) = \sum_{E' \subseteq E} (-1)^{|E|-|E'|} \times \mathbb{I}[E' \text{ connects } x, y, z]
\]

\[d(G) = 4\]
**Example calculation**

**MSF path vertex**

\[ d_{MSF}(G|\pi) = \sum_{E' \subseteq E} (-1)^{|E|-|E'|} \mathbb{I}[E' \text{ connects } x, y] \times \mathbb{I}[\gamma_{MST}(G'_{E'}|\pi'_{E'}) \text{ passes through } z] \]

\[ d_{MSF}(G|\pi) = 1 \]

---

**Diagram**

- Example calculations for different paths and edges connecting vertices:
  - Path through vertices 1, 2, 3, 4, 5, 6, 1, 2.

- Two paths highlighted in red for demonstration:
  - Path through vertices 1, 2, 3.
  - Path through vertices 4, 5.

- Results for each path:
  - Path through vertices 1, 2, 3: +1
  - Path through vertices 4, 5: 0

---

**Table**

<table>
<thead>
<tr>
<th>Path Configuration</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 3</td>
<td>+1</td>
</tr>
<tr>
<td>4, 5</td>
<td>0</td>
</tr>
</tbody>
</table>
Example calculation

**MSF path vertex**

\[
d_{\text{MSF}}(G|\pi) = \sum_{E' \subseteq E} (-1)^{|E|-|E'|} \mathbb{I}[E' \text{ connects } x, y] \\
\times \mathbb{I}[\gamma_{\text{MST}}(G'_E|\pi'_E) \text{ passes through } z]
\]

\[d_{\text{MSF}}(G|\pi') = 0\]
Lattice to continuum

Sum over graphs of same topology $\to$ continuum expansion $d(G), d_{MSF}(G | \pi)$ invariant under insertion of paths of edges

Proper weight for orderings?

Toy Example:

Only need keep max. cost edge (barrier) along any path

$$\Pr[L_1 < L_2 \leq p_0] = \int_0^{p_0} dp_2 \int_0^{p_2} dp_1 \frac{d}{dp_1} \frac{d}{dp_2} P_G(p_1, p_2) = \frac{N_2}{N_1 + N_2} p_0^{N_1 + N_2}$$

Generalization straightforward...
Continuum theory

• Generating functions of random walk = massive propagator for scalar field
• Can neglect excluded volume constraint for RG
• Interpret mass on each propagator as MST barrier for given path
  • Higher barrier \( p \rightarrow \) formed later as Kruskal is run \( \rightarrow \) smaller value of \( t \)

\[
\mathcal{O}_{\text{MSF}}(\pi, t_0) = \int \cdots \int_{\infty > t_{\pi(1)} \geq t_{\pi(2)} \geq \cdots \geq t_{\pi(n)} \geq t_0} \prod_{\epsilon \in E} dt_{\epsilon} \frac{d}{dt_{\epsilon}} \text{ acting on Feynman diagram of } \varphi^3 \text{ theory}
\]
Continuum theory

- Generating functions of random walk = massive propagator for scalar field
- Can neglect excluded volume constraint for RG
- Interpret mass on each propagator as MST barrier for given path
- Higher barrier $p \rightarrow$ formed later as Kruskal is run $\rightarrow$ smaller value of $t$

$$\mathcal{O}_{\text{MSF}}(\pi, t_0) = \int \cdots \int_{\infty > t_{\pi(1)} \geq t_{\pi(2)} \geq \cdots \geq t_{\pi(n)} \geq t_0} \prod_{\epsilon \in E} dt_{\epsilon} \frac{d}{dt_{\epsilon}}$$

Diagrammatic weights unchanged

acting on Feynman diagram of $\varphi^3$ theory
To summarize

• To generate the Feynman graph expansion for the MSF path vertex function,
  • Start with graphs for the mass insertion in a massive cubic theory;
  • For each graph and each ordering of edge masses,
    • compute $d_{\text{MSF}}(G|\pi)$ for that ordering,
    • assign each edge a unique mass and act with $O_{\text{MSF}}(t_0)$ for that ordering and
    • sum over orderings and perform integration over loop momenta of Feynman diagram.

$$\tilde{C}_{x,y}(p_0) \rightarrow \sum_{G} \sum_{\pi' \in S_{|E|}} d_{\text{MSF}}(G|\pi') O_{\text{MSF}}(\pi', t_0) I(G, \{t_\epsilon\})$$
“Renormalizability”

- The (possible) problem:

\[ \Gamma^{PV} = \text{\begin{tikzpicture}[baseline=-0.5ex] \node (a) at (0,0) {}; \node (b) at (0.5,0) {}; \node (c) at (1,0) {}; \draw (a) -- (b); \draw (b) -- (c); \end{tikzpicture}} + \text{\begin{tikzpicture}[baseline=-0.5ex] \node (a) at (0,0) {}; \node (b) at (0.5,0) {}; \node (c) at (1,0) {}; \draw (a) -- (b) -- (c); \end{tikzpicture}} + \text{\begin{tikzpicture}[baseline=-0.5ex] \node (a) at (0,0) {}; \node (b) at (0.5,0) {}; \node (c) at (1,0) {}; \draw (a) -- (b) -- (c); \node (d) at (1.5,0) {}; \draw (c) -- (d); \end{tikzpicture}} + \ldots \rightarrow \text{\begin{tikzpicture}[baseline=-0.5ex] \node (a) at (0,0) {}; \node (b) at (0.5,0) {}; \node (c) at (1,0) {}; \draw (a) -- (b) -- (c); \end{tikzpicture}} \]
“Renormalizability”

• The (possible) problem:

\[ \Gamma^{PV} = \gamma + \Delta + \Delta + \Delta + \ldots \]

• Need contribution from any subgraph to be independent of costs outside
“Renormalizability”

- The (possible) problem:

\[ \Gamma^{PV} = \sum \text{Terms} \rightarrow \text{Result} \]

- Need contribution from any subgraph to be independent of costs outside
- *Not true* for lattice MST or expansion!
“Renormalizability”

• The (possible) problem:

\[ \Gamma^{PV} = \quad + \quad + \quad + \quad + \cdots \quad \rightarrow \quad \]

• Need contribution from any subgraph to be independent of costs outside

• *Not true* for lattice MST or expansion!

• RG simplifications:
“Renormalizability”

- The (possible) problem:
  \[ \Gamma^{pv} = \ldots \]

- Need contribution from any subgraph to be independent of costs outside
  - *Not true* for lattice MST or expansion!

- RG simplifications:
  - Only keep UV-divergent terms = positive powers of $\Lambda$
“Renormalizability”

• The (possible) problem:

\[ \Gamma^{PV} = \ldots + \ldots + \ldots + \ldots + \ldots \]

• Need contribution from any subgraph to be independent of costs outside

• *Not true* for lattice MST or expansion!

• RG simplifications:
  • Only keep UV-divergent terms = positive powers of \( \Lambda \)
  • 2- and 3-pt. subgraphs only
“Renormalizability”

• The (possible) problem:

\[ \Gamma^{PV} = \Gamma + \Gamma + \Gamma + \Gamma + \ldots \rightarrow \land \]

• Need contribution from any subgraph to be independent of costs outside
  • *Not true* for lattice MST or expansion!

• RG simplifications:
  • Only keep UV-divergent terms = positive powers of \( \Lambda \)
  • 2- and 3-pt. subgraphs only
“Renormalizability”

• $O_{MSF}$ factorizes

• Only UV-divergent contributions to any diagram come from a *subset* of all orderings
  • Short paths form first $\rightarrow$ High-momentum propagators get higher masses

• $d_{MSF}$ factorizes for these orderings (*not* in general)
  • Two-point diagram gives barrier mass insertion
(RG-relevant) **Factorization of** $d_{\text{MSF}}$

\[
d_{\text{MSF}} \left( g_1, g_2, \ldots, g_z, \ldots, g_n \right) = d_{\text{MSF}} \left( x, g_z, y \right) \times \prod_{i=1}^{n} d \left( g_i \right)
\]

(1PI)

**Percolation**

\[
d_{\text{MSF}} \left( x, g_z, y \right) = d \left( x, g_z, y \right) \times d_{\text{MSF}} \left( x, y \right)
\]

**Path vertex:** \( d_{\text{MSF}} \left( x, g_z, y \right) = d_{\text{MSF}} \left( x, g_z, y \right) \times d_{\text{MSF}} \left( t_H, x, y \right) \)

**Barrier:** \( d_{\text{MSF}} \left( x, g_z, y \right) = d \left( x, g_z, y \right) \times d_{\text{MSF}} \left( t_H, x, y \right) \)
1 loop RG

\begin{align*}
  d_2 &= -1 \\
  d_3 &= -2 \\
  d_{MSF}(\triangle | t_1, t_2 > t_3) &= 0 \\
  \text{else} &= -1 \\
  \Rightarrow d_{PV} &= -\frac{2}{3}
\end{align*}

\begin{align*}
  \beta(u_0) &= -\frac{\varepsilon}{2} u_0 + \left( \frac{d_2}{4} - d_3 \right) u_0^3 + O(u_0^5, \varepsilon u_0^3) \\
  \gamma_\phi(u_0) &= \frac{d_2}{6} u_0^2 + O(u_0^4, \varepsilon u_0^2) \\
  \gamma_{PV}(u_0) &= \left( d_{PV} - \frac{d_2}{6} \right) u_0^2 + O(u_0^4, \varepsilon u_0^2)
\end{align*}

\begin{align*}
  D_{MSF} &= 2 - \frac{(6d_{PV} - d_2)}{3(d_2 - 4d_3)} \varepsilon = 2 - \frac{\varepsilon}{7} + O(\varepsilon^2)
\end{align*}

singly connected bonds: \( D_{perc} = \nu_{perc}^{-1} = 2 - \frac{5\varepsilon}{21} + O(\varepsilon^2) \)
Future directions

• Behavior for $p > p_c$

• Application to other problems?
  • Loop-erased random walk from self-avoiding walk
  • Two-replica formulation of $\pm J$ EA spin glass
    • Metastable states???
    • Ground state enumeration?

• Perturbation away from strong disorder
  • Failure of greedy algorithms related to, e.g., RSB instability?
Summary

• MST problem has $d_c = 6$, not 8
• $D_{\text{MST}} = 2 - \frac{\varepsilon}{7} + \mathcal{O}(\varepsilon^2)$
• Treatment of quenched disorder in strong limit without (explicit?) replicas, cavity method, FRG, supersymmetry,...
  • Nonlocality → perturbation series not generated by (local) action
• Simplification within RG framework

Part II: TSJ and N. Read, arXiv:0909.5343