Weighted Monte-Carlo sampling of Feynman graphs in ϕ^4-theory

Paul-Hermann Balduf, U of Waterloo & Perimeter Institute

Theory Canada, May 24, 2024

based on arXiv 2403.16217 with Kimia Shaban

Slides, references, data set etc. available from paulbalduf.com/research

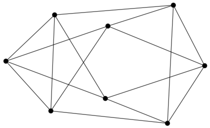
Background

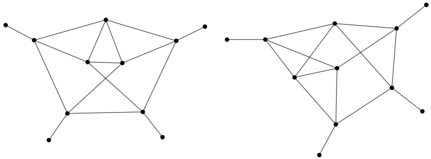
- Perturbative quantum field theory in flat (Euclidean) $D = 4 - 2\epsilon$ spacetime.
- Massless bosonic ϕ^4-theory

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{\lambda}{4!}\left(\phi^2\right)^2.$$

$\Rightarrow$ Feynman graphs have 1 type of edge, 1 type of 4-valent vertex.

- We want to compute the beta function of this theory.





The beta function in ϕ^4-theory

- Describes change of the coupling when *all* energy scales are increased simultaneously. Set $\ell := \ln \frac{p^2}{\mu^2}$, where μ arbirary reference. *Callan-Symanzik equation* for renormalized Green functions $\mathcal{G}(\alpha, \ell)$ Callan 1970; Symanzik 1970:

$$\partial_\ell \mathcal{G}(\alpha, \ell) = \Big(\gamma(\alpha) + \beta(\alpha)\alpha\partial_\alpha\Big)\mathcal{G}(\alpha, \ell).$$

- Related to running coupling $\alpha(\mu)$ and coupling counterterm Z_α via $\beta(\alpha) = \alpha\mu\frac{\partial}{\partial\mu}\ln \alpha(\mu, \alpha_0) = \frac{-\epsilon}{\partial_\alpha \ln(\alpha Z_\alpha(\alpha,\epsilon))}$.
- Computed perturbatively from sum of vertex-type L-loop Feynman integrals $\mathcal{F}(G)$,

$$\beta(\alpha) = -2 \sum_{\text{4-valent graphs } G} (-\alpha)^{L+1} \frac{\partial_\ell \mathcal{F}(G)}{|\mathrm{Aut}(G)|}.$$

Periods in ϕ^4-theory

- We consider only *primitive* (=no subdivergences) graphs G in $D = 4 - 2\epsilon$.

$$\mathcal{F}(G) = \text{const} \cdot \left(\frac{1}{\epsilon} \frac{\mathcal{P}(G)}{L} - \mathcal{P}(G) \cdot \ell + \ell\text{-independent terms} + \mathcal{O}(\epsilon) \right)$$

- First Symanzik polynomial ψ_G. Nontrivial part of integral is the *period* [Broadhurst and Kreimer 1995; Schnetz 2010]

$$\mathcal{P}(G) = \left(\prod_{e \in E_G} \int_0^\infty \mathrm{d}a_e \right) \delta\left(1 - \sum_{e=1}^{|E_G|} a_e \right) \frac{1}{\psi_G^2(\{a_e\})} \in \mathbb{R}.$$

- With the period, $\partial_\ell \mathcal{F}(G) = \mathcal{P}(G)$. The *primitive* contribution to the beta function is

$$\beta^{\text{prim}}(\alpha) = 2 \sum_{\text{primitive 4-valent } G} (-\alpha)^{L+1} \frac{\mathcal{P}(G)}{|\text{Aut}(G)|}.$$

Computing the primitive beta function

- Periods can be quickly ($\sim$ 1h/graph) computed numerically with new algorithm up to $L \approx 16$ loops [Borinsky 2023; Borinsky, Munch, and Tellander 2023]
- Can exploit various symmetries, only a subset is truly independent [Schnetz 2010; Panzer 2022; Hu et al. 2022]
- 2 Problems (see my talk on May 28):
  - Number of graphs grows factorially, 750k at 13 loops, 950M at 16 loops $\Rightarrow$ impossible to compute all of them, need (random) sample, *Monte Carlo* algorithm.
  - Standard deviation of distribution is large, $\sigma(\mathcal{P}) \approx \langle \mathcal{P} \rangle \Rightarrow$ sample has large statistical uncertainty, at sample size n

$$\Delta_{\text{samp}} \mathcal{P} = \frac{1}{\sqrt{n}} \sigma(\mathcal{P}), \quad \Rightarrow \quad \frac{\Delta_{\text{samp}} \mathcal{P}}{\mathcal{P}} \approx \frac{1}{\sqrt{n}}.$$

    E.g. for 3 significant digits ($\Delta_{\text{samp}} \leq 0.1\%$) we need sample size $n \approx 10^6$.
- Solution: Importance sampling of periods.

Importance sampling for periods

- Idea of importance sampling: If we know a function $\bar{\mathcal{P}}$ which approximates the period *and* $\bar{\mathcal{P}}$ is fast to compute, then:
  1. Evaluate $\langle \bar{\mathcal{P}} \rangle$ in a large sample of size $N_s \cdot n$.
  2. Generate a smaller random sample S of n graphs weighted proportional to $\bar{\mathcal{P}}$. Evaluate $\left\langle \frac{\mathcal{P}}{\bar{\mathcal{P}}} \right\rangle_S$ in this sample.
  3. Law of conditional probability:

$$\langle \mathcal{P} \rangle = \underbrace{\langle \bar{\mathcal{P}} \rangle}_{\text{large sample, fast}} \cdot \underbrace{\left\langle \frac{\mathcal{P}}{\bar{\mathcal{P}}} \right\rangle_S}_{\text{small sample, slow}}$$

- Total error is small if *simultaneously* $\delta := \sigma\left(\frac{\mathcal{P}}{\bar{\mathcal{P}}}\right)$ is small (high accuracy of approximation) and $N_s \gg 1$ (approximation faster than numerical integration)

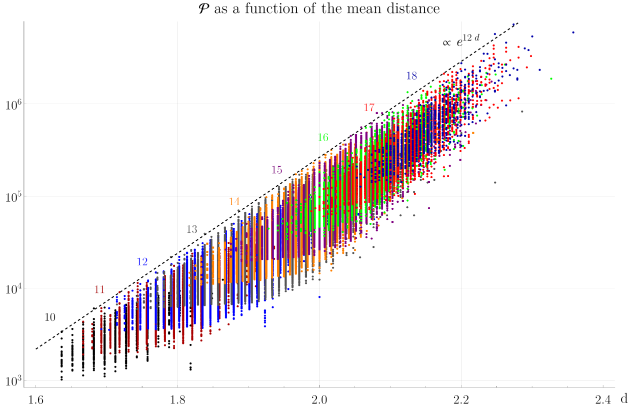
Approximation functions

We examined ≈ 150 different graph-theoretical quantities empirically, for a data set [Balduf 2023, freely available] of $\approx 1.5M$ periods with $L \leq 18$.

Recall that at fixed L, all graphs have the same number of edges and vertices and are 4-regular.

Average vertex distance

- Count number of edges between all pairs of vertices, take average.
- Relatively fast to count, clearly correlated, but low accuracy $\delta \approx 30\%$.



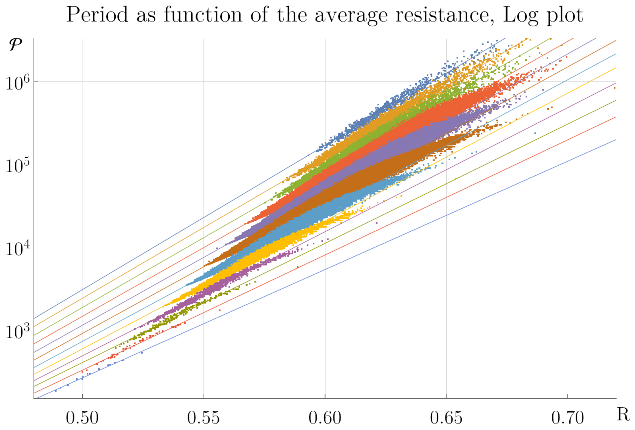
Average resistance (Kirchhoff index)

- Assign unit electrical resistance to every edge. Resistance r_{v_i,v_j} between vertices v_i and v_j.
- *Kirchhoff index* = average resistance

$$R(G) := \frac{2}{|V_G|(|V_G|-1)} \sum_{v_1<v_2\in V_G} r_{v_1,v_2}.$$

- Extremely fast to compute due to matrix linear algebra operations ($\sim 100\mu$s per graph)

Average resistance



(Colors = loop orders 5 . . . 18)

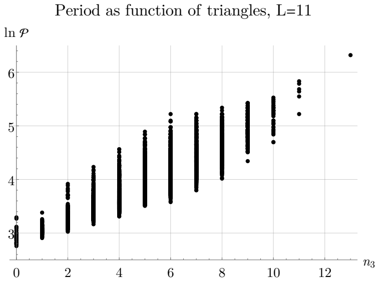
- Resistance approximation reaches $\sigma\left(\frac{\mathcal{P}}{\bar{\mathcal{P}}}\right) =: \delta \approx 5\%$.

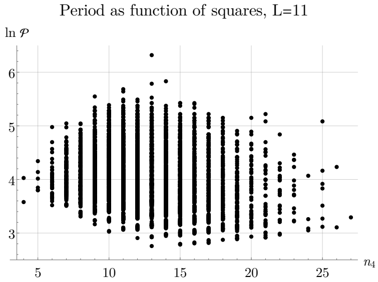
Cycles

- $n_j(G)$ the total number of cycles of length $j \in \mathbb{N}$ contained in a graph G.
- $n_1 = n_2 = 0, \quad n_3 =$ number of triangles, $n_4 =$number of squares, ...

Cycles

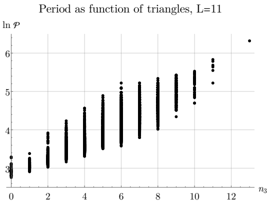
- $n_j(G)$ the total number of cycles of length $j \in \mathbb{N}$ contained in a graph G.
- $n_1 = n_2 = 0, \quad n_3 =$ number of triangles, n_4=number of squares, ...
- Triangles are a mediocre approximation, $\mathcal{P}$ almost uncorrelated with n_j for $j \geq 4$.

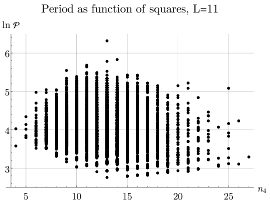


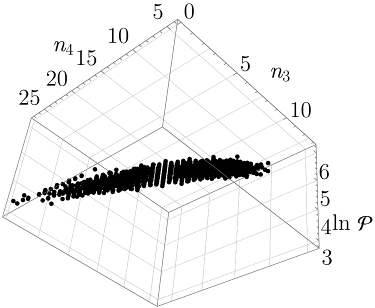


Cycles

- Triangles are a mediocre approximation, $\mathcal{P}$ almost uncorrelated with n_j for $j \geq 4$.
- But: n_3 and n_4 together are good!
- $\Rightarrow$ Use multi-linear function.







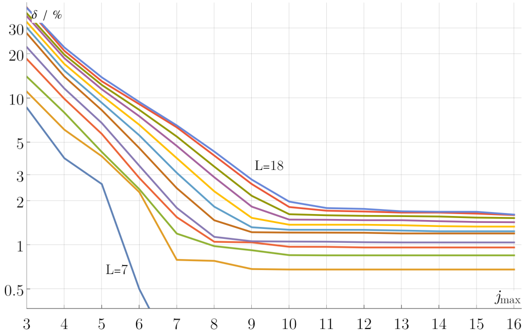
Cycles

- Use multi-linear function of j-cycle count n_j,

$$\ln \bar{\mathcal{P}} := f_0 + \sum_{j=3}^{j_{\max}} f_j \frac{2j \cdot n_j}{3^j}.$$

- Approximation gets better with increasing $j_{\max}$. Saturated at $j \approx 10$, accuracy $\delta \approx 2\%$.

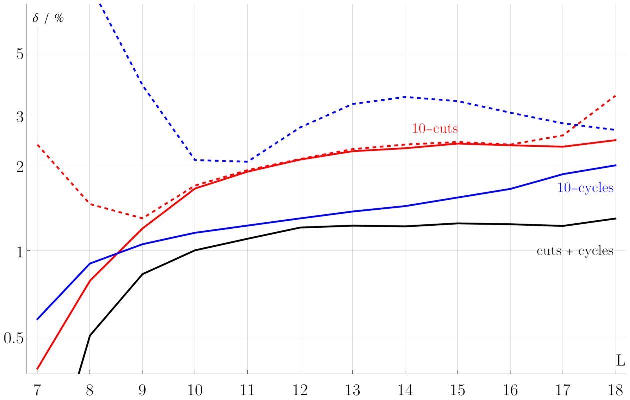
Relative standard deviation as function of $j_{\max}$, Log plot



Cuts

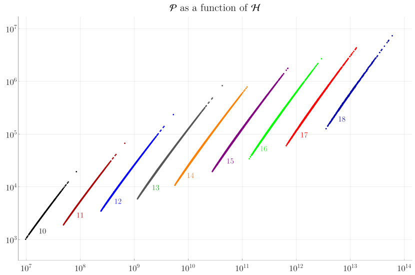
- A (minimal) cut $C \subseteq E_G$ separates the grap into exactly 2 connected components. c_j number of j-edge cuts. Consider multi-linear model of $\ln(c_j)$.
- Can combine cuts and cycles. Reach $\delta \approx 1.2\%$ for all loop orders.

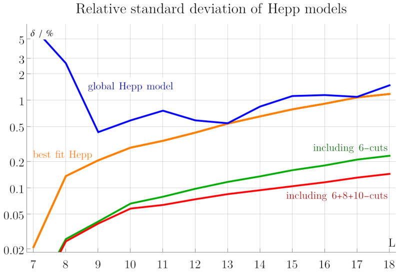
Relative standard deviation of Cut + Cycle models, All graphs



Hepp bound.

- *Hepp bound* $\mathcal{H}(G)$ [Hepp 1966; Panzer 2022] arises from "tropicalization" of period integral upon sector decomposition.
- Strongly correlated with period. Polynomial function, combined with edge-cuts $\ln(c_j)$, gives $\delta \approx 0.2\%$.
- Computation requires iteration over *all* subgraphs (and/or caching).

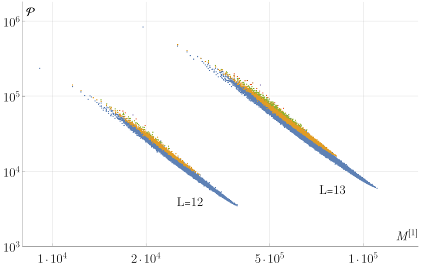




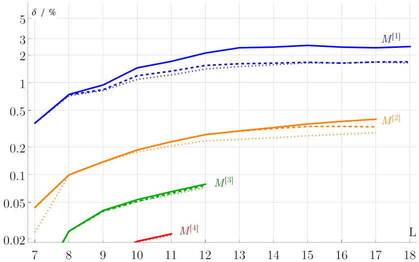
Martin invariant

- k^{th} order *Martin invariant* $M^{[k]}$ [Panzer and Yeats 2023] is derivative of $O(N)$ symmetry factor at $N=-2$ for a graph where every edge is replaced by k parallel edges.
- Linear function of $\ln M^{[1]}$ gives $\delta \approx 4\%$, higher $M^{[k]}$ are much better. k^{th}-order polynomial of $M^{[k]}$ can get very accurate when cobined with cuts $\ln(c_j)$, reach $\delta \ll 0.1\%$.
- Like Hepp, requires recurrence over decompositions and caching.

Period as a function of Martin invariant, Double Log plot



Cubic Martin+Cut model



Machine learning models

- So far: Hand-picked, physically inspired quantities.
- Linear regression of *all* quantities simultaneously gives $\delta \approx 0.1\%$. Quadratic even better.
  But: High computational cost.

Work by **Kimia Shaban**

Machine learning models

Work by **Kimia Shaban**

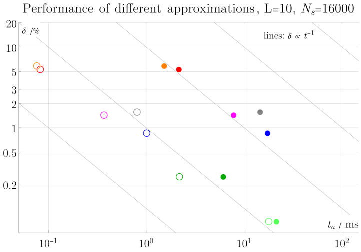
- So far: Hand-picked, physically inspired quantities.
- Linear regression of *all* quantities simultaneously gives $\delta \approx 0.1\%$. Quadratic even better.
  But: High computational cost.
- Trained artificial neural networks (graph convolutional network, graphSAGE) using *just* the graph as input.
- Not accurate/reproducible enough for Monte Carlo sampling, but very fast once trained.
- Machine learning challenge: The graph *completely* determines the period (perfect accuracy is possible for a clever enough model). Data set freely available, see paulbalduf.com/research.

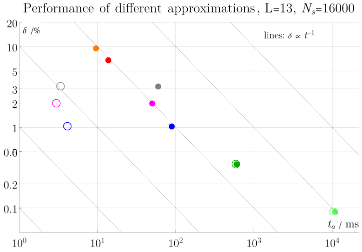
Speed vs. accuracy

- Recall: Sampling accuracy scales as $\frac{1}{\sqrt{n}}\delta$, where $n \propto \frac{1}{t_a}$. $\Rightarrow$ equivalent approximations on lines $\delta \propto t_a^{-\frac{1}{2}}$.

Speed vs. accuracy

- Recall: Sampling accuracy scales as $\frac{1}{\sqrt{n}}\delta$, where $n \propto \frac{1}{t_a}$. $\Rightarrow$ equivalent approximations on lines $\delta \propto t_a^{-\frac{1}{2}}$.
- Most models lie on $\delta \propto t_a^{-1}$. I.e. the more accurate models are "more than worth" their extra time. Still, Hepp and Martin are too slow for $L > 14$.
- $\Rightarrow$ Cut & Cycle model is the most useful one for Monte-Carlo sampling.





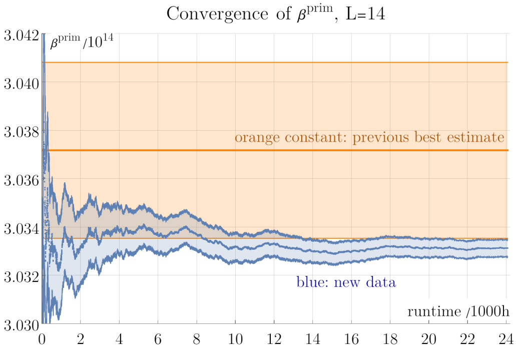
(filled: 1 thread, hollow: 20 threads)

Weighted sampling implementation details

- Generate random graphs, check for primitivity, compute "cuts & cycles" prediction $\bar{\mathcal{P}}$
- Metropolis-Hastings sampling algorithm: Have graph G_1. Generate $x \in [0,1]$ uniformly. If $\frac{\bar{\mathcal{P}}(G_2)}{\bar{\mathcal{P}}(G_1)} > x$, accept G_2, else keep G_1.
- In our case, sampling factor $N_s \approx 5000$ approximations per integration was ideal to balance accuracy of sampling $\left\langle \frac{\mathcal{P}}{\bar{\mathcal{P}}} \right\rangle$ vs $\langle \bar{\mathcal{P}} \rangle$ sampling.
- Distribute 100 threads on prediction and integration dynamically.

Example results: Primitive beta function for $L = 14$

- Reached 120ppm standard deviation after 24k CPU core h ($<$ 2 weeks walltime).



- Previous work with uniform random sampling took 400k CPU core h for 1063ppm.
- $\Rightarrow$ Weighted sampling is $\approx 1000\times$ faster than uniform random sampling, or reaches $\approx 35\times$ the accuracy at the same runtime.

Conclusion

- Feynman integrals are correlated to various properties of the graph.
- Some correlations are conceptually interesting (e.g. resistance). Often, $\ln(\mathcal{P})$ is a multi-linear function.
- To approximate period $\mathcal{P}$, we can reach approximation uncertainty $\delta < 2\%$ easily, $\delta \approx 0.1\%$ with some effort.
- Weighted sampling reduces the time for numerically computing the primitive beta function by roughly a factor 10^3, but only for $L \geq 13$ loops (otherwise, just compute all graphs).

Thank you!

References I

Balduf, Paul-Hermann (2023). "Statistics of Feynman Amplitudes in ϕ^4-Theory". In: *Journal of High Energy Physics* 2023.11, p. 160. DOI: 10.1007/JHEP11(2023)160.

Borinsky, Michael (2023). "Tropical Monte Carlo Quadrature for Feynman Integrals". In: *Annales de l'Institut Henri Poincaré D* 10.4, pp. 635–685. DOI: 10.4171/AIHPD/158. arXiv: 2008.12310. arXiv: 2008.12310.

Borinsky, Michael, Henrik J. Munch, and Felix Tellander (2023). *Tropical Feynman Integration in the Minkowski Regime*. DOI: 10.48550/arXiv.2302.08955. arXiv: 2302.08955 [hep-ph, physics:hep-th, physics:math-ph]. arXiv: 2302.08955. preprint.

Broadhurst, D. J. and D. Kreimer (1995). "Knots and Numbers in ϕ^4 Theory to 7 Loops and Beyond". In: *International Journal of Modern Physics C* 06.04, pp. 519–524. DOI: 10.1142/S012918319500037X. arXiv: hep-ph/9504352. arXiv: hep-ph/9504352.

Callan, Curtis G. (1970). "Broken Scale Invariance in Scalar Field Theory". In: *Physical Review D* 2.8, pp. 1541–1547. DOI: 10.1103/PhysRevD.2.1541.

Hepp, Klaus (1966). "Proof of the Bogoliubov-Parasiuk Theorem on Renormalization". In: *Communications in Mathematical Physics* 2.1, pp. 301–326. DOI: 10.1007/BF01773358.

Hu, Simone et al. (2022). "Further Investigations into the Graph Theory of ϕ^4-Periods and the c_2 Invariant". In: *Annales de l'Institut Henri Poincaré D* 9.3, pp. 473–524. DOI: 10.4171/AIHPD/123.

Panzer, Erik (2022). "Hepp's Bound for Feynman Graphs and Matroids". In: *Annales de l'Institut Henri Poincaré D* 10.1, pp. 31–119. DOI: 10.4171/aihpd/126.

References II

Panzer, Erik and Karen Yeats (2023). *Feynman Symmetries of the Martin and c_2 Invariants of Regular Graphs*. DOI: 10.48550/arXiv.2304.05299. arXiv: 2304.05299 [hep-th, physics:math-ph]. arXiv: 2304.05299. preprint.

Schnetz, Oliver (2010). "Quantum Periods: A Census of ϕ^4-Transcendentals". In: *Commun.Num.Theor.Phys.* 4, pp. 1–48. arXiv: 0801.2856. arXiv: 0801.2856.

Symanzik, Kurt (1970). "Small Distance Behaviour in Field Theory and Power Counting". In: *Communications in Mathematical Physics* 18.3, pp. 227–246. DOI: 10.1007/BF01649434.