Thermodynamic Semirings and Information Algebras

Matilde Marcolli

Ma148b: Algebraic and Categorical Aspects of Information
Caltech, Winter 2025

Based on:

- M. Marcolli, R. Thorngren, *Thermodynamic semirings*, J. Noncommut. Geom. 8 (2014), no. 2, 337–392
- N.Combe, Yu.I.Manin, M.Marcolli, *Quantum Operads*, work in progress

Min-Plus Algebra (Tropical Semiring)

min-plus (or tropical) semiring $\mathbb{T} = \mathbb{R} \cup \{\infty\}$

- operations $\oplus$ and $\odot$

$$x \oplus y = \min\{x, y\} \quad \text{with identity } \infty$$

$$x \odot y = x + y \quad \text{with identity } 0$$

- operations $\oplus$ and $\odot$ satisfy:
  - associativity
  - commutativity
  - left/right identity
  - distributivity of product $\odot$ over sum $\oplus$

Note: can work equivalently with $(\mathbb{R} \cup \{\infty\}, \min, +)$ or with $(\mathbb{R}_+, \max, *)$ isomorphic under $-\log$ map

Convexity in characteristic one semirings

- on $K = (\mathbb{R}_+, \max, *)$ partial ordering $\leq$ by $x \leq y \Leftrightarrow x \oplus y = y = \max\{x, y\} = y$
- more generally K commutative characteristic one semifield (ie with $\odot$-multiplicative inverses and where $1 \oplus 1 = 1$ idempotent)
- idempotent property and distributive property imply characteristic one Frobenius automorphism:

$$(x \oplus y)^{\odot n} = x^{\odot n} \oplus y^{\odot n}$$

- such K is $(\mathbb{R}_+, \max, *)$-module through $(x, t) \mapsto x^t$
- function $f : X \to K$ with X a convex subset of topological $(\mathbb{R}_+, \max, *)$-module is convex:

$$f(tx_1 + (1-t)x_2) \leq f(x_1)^t f(x_2)^{1-t}$$

- usual definition of convexity when r.h.s. in $(\mathbb{R} \cup \{\infty\}, \min, +)$

Legendre transform

- *epigraph* $\mathrm{epi} f = \{(\alpha, r) \in X \times K | f(\alpha) \leqslant r\}$
- f is convex iff the epigraph $\mathrm{epi} f$ is convex and f is closed iff $\mathrm{epi} f$ is closed
- for $X \subseteq K$, Legendre transform of $f : X \to K$

$$f^*(x) = \sum_{\alpha \in X} \frac{x^\alpha}{f(\alpha)}$$

- on $(\mathbb{R} \cup \{\infty\}, \min, +)$ this is usual Legendre transform

$$f^*(x) = \sup_{\alpha \in X} (\alpha x - f(\alpha)),$$

- Legendre transform of f is closed and convex (by epigraph)
- (Fenchel-Moreau) for $f : X \to K$ with $X \subset \mathbb{R}_{\geqslant 0}$
  1. f^{**} is closed and convex and bounded by f
  2. $f^{**} = f$ iff f is closed and convex

Thermodynamic semirings $\mathbb{T}_{\beta,S} = (\mathbb{R} \cup \{\infty\}, \oplus_{\beta,S}, \odot)$

- deformation of the tropical addition $\oplus_{\beta,S}$

$$x \oplus_{\beta,S} y = \min_p \{px + (1-p)y - \frac{1}{\beta} S(p)\}$$

β thermodynamic inverse temperature parameter
$S(p) = S(p, 1-p)$ binary information measure, $p \in [0,1]$

- for $\beta \to \infty$ (zero temperature) recovers unperturbed idempotent addition $\oplus$

- multiplication $\odot = +$ is undeformed

- for $S =$ Shannon entropy considered first in relation to $\mathbb{F}_1$-geometry in
  - A. Connes, C. Consani, *From monoids to hyperstructures: in search of an absolute arithmetic*, arXiv:1006.4810

Khinchin axioms $\mathrm{Sh}(p) = -C(p\log p + (1-p)\log(1-p))$

- Axiomatic characterization of Shannon entropy $S(p) = \mathrm{Sh}(p)$
  1. symmetry $S(p) = S(1-p)$
  2. minima $S(0) = S(1) = 0$
  3. extensivity
     $S(pq) + (1-pq)S(p(1-q)/(1-pq)) = S(p) + pS(q)$
- correspond to algebraic properties of semiring $\mathbb{T}_{\beta,S}$
  1. commutativity of $\oplus_{\beta,S}$
  2. left and right identity for $\oplus_{\beta,S}$
  3. associativity of $\oplus_{\beta,S}$

$\Rightarrow \mathbb{T}_{\beta,S}$ commutative, unital, associative iff $S(p) = \mathrm{Sh}(p)$

case of associativity:

$$x \oplus_{\beta,S} (y \oplus_{\beta,S} z) = x \oplus_{\beta,S} \min_p (py + (1-p)z - \frac{1}{\beta} S(p))$$

$$= \min_q (qx + (1-q) \min_p (py + (1-p)z - \frac{1}{\beta} S(p)) - \frac{1}{\beta} S(q))$$

$$= \min_{p,q} (qx + p(1-q)y + (1-q)(1-p)z - \frac{1}{\beta}(S(q) + (1-q)S(p)))$$

$$= \min_{p_1+p_2+p_3=1} (p_1 x + p_2 y + p_3 z - \frac{1}{\beta}(S(p_1) + (1-p_1)S(\frac{p_2}{1-p_1})))$$

while

$$(x \oplus_{\beta,S} y) \oplus_{\beta,S} z = \min_p (px + (1-p)y - \frac{1}{\beta} S(p)) \oplus_{\beta,S} z$$

$$= \min_{p,q} (pqx + q(1-p)y + (1-q)z - \frac{1}{\beta}(qS(p) + S(q))$$

$$= \min_{p_1+p_2+p_3=1} (p_1 x + p_2 y + p_3 z - \frac{1}{\beta}(S(p_1+p_2) + (p_1+p_2)S(\frac{p_1}{p_1+p_2})).$$

Khinchin axioms *n*-ary form

Given S as above, define $S_n : \Delta_{n-1} \to \mathbb{R}_{\geqslant 0}$ by

$$S_n(p_1, \ldots, p_n) = \sum_{1 \leqslant j \leqslant n-1} (1 - \sum_{1 \leqslant i < j} p_i) S(\frac{p_j}{1 - \sum_{1 \leqslant i < j} p_i}).$$

Then Khinchin axioms:

1. (Continuity) $S(p_1, \ldots, p_n)$ continuous in $(p_1, \ldots, p_n) \in \Delta_n$ simplex
2. (Maximality) $S(p_1, \ldots, p_n)$ maximum at the uniform $p_i = 1/n$
3. (Additivity/Extensivity) $p_i = \sum_{j=1}^{m_i} p_{ij}$ then

$$S(p_{11}, \ldots, p_{nm_n}) = S(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i S(\frac{p_{i1}}{p_i}, \ldots, \frac{p_{im_i}}{p_i});$$

4. (Expandability) Δ_n face in Δ_{n+1}

$$S(p_1, \ldots, p_n, 0) = S(p_1, \ldots, p_n)$$

extensivity axiom

- $(J_k)_{1\leqslant k\leqslant m}$ be a partition of $\{p_1,\ldots,p_n\}$
- $S_n(p_1,\ldots,p_n)$ defined in terms of binary $S(p)$ as

$$S_n(p_1,\ldots,p_n)=\sum_{1\leqslant j\leqslant n-1}(1-\sum_{1\leqslant i<j}p_i)S(\frac{p_j}{1-\sum_{1\leqslant i<j}p_i})$$

- then have

$$S_n(p_1,\ldots,p_n)=S_m(q_1,\ldots,q_m)+\sum_{1\leqslant k\leqslant m}q_k\, S_{|J_k|}(J_k/q_k),$$

where $q_k=\sum_{p\in J_k}p$, so J_k/q_k is a $|J_k|$-ary probability distribution

Shannon entropy case:

$$x \oplus_{\beta,\mathrm{Sh}} y = \min_p \{ px + (1-p)y - \frac{1}{\beta}\mathrm{Sh}(p) \}$$

equivalent form of $\oplus_{\beta,\mathrm{Sh}}$

$$x \oplus_{\beta,\mathrm{Sh}} y = -\beta^{-1} \log \left(e^{-\beta x} + e^{-\beta y} \right)$$

for $T = 1/\beta$ (temperature parameter) and on tropical semiring $(\mathbb{R} \cup \{\infty\}, \min, +)$

$$x \oplus_{T,\mathrm{Sh}} y = -T \log(e^{-x/T} + e^{-y/T})$$

or with multiplicative notation $(\mathbb{R}_+, \max, *)$

$$x \oplus_{T,\mathrm{Sh}} y = (x^{1/T} + y^{1/T})^T$$

- $F(p) = px + (1-p)y + T(p\log p + (1-p)\log(1-p))$
- $\partial F(p) = x - y + T(\log p - \log(1-p))$
- $\partial F(p) = 0$ when $T\log p + x - y = T\log(1-p)$

$$p_{min} = \frac{1}{1 + e^{\frac{x}{T} - \frac{y}{T}}} = \frac{e^{-x/T}}{e^{-x/T} + e^{-y/T}}$$

- so get at minimum in p

$$F(p_{min}) = -T\log\left(e^{-x/T} + e^{-y/T}\right)$$

Expressing Shannon thermodynamic addition as

$$x \oplus_{T,\mathrm{Sh}} y = (x^{1/T} + y^{1/T})^T$$

leads to relation with Maslov dequantization and tropical geometry

Relation to tropical geometry

- tropical polynomial $p : \mathbb{R}^n \to \mathbb{R}$ piecewise linear

$$p(x_1, \ldots, x_n) = \oplus_{j=1}^m a_j \odot x_1^{k_{j1}} \odot \cdots \odot x_n^{k_{jn}} =$$

$$\min\{a_1 + k_{11}x_1 + \cdots + k_{1n}x_n, a_2 + k_{21}x_1 + \cdots + k_{2n}x_n, \cdots, a_m + k_{m1}x_1 + \cdots + k_{mn}x_n\}.$$

tropical hypersurface where tropical polynomial non-differentiable

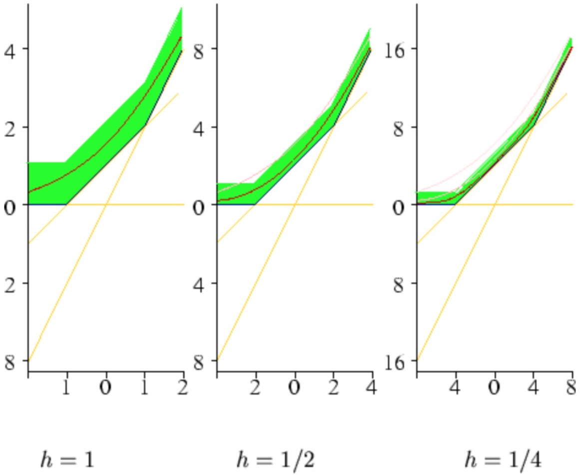
- Entropical geometry: thermodynamic deformations of $\mathbb{T}$

$$p_{\beta,\mathcal{S}}(x_1, \ldots, x_n) = \oplus_{\beta,S,j} a_j \odot x_1^{k_{j1}} \odot \cdots \odot x_n^{k_{jn}} =$$

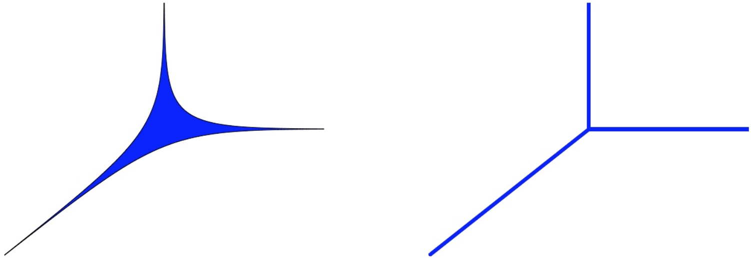
$$\min_{p=(p_j)} \{\sum_j p_j(a_j + k_{j1}x_1 + \cdots + k_{jn}x_n) - \frac{1}{\beta} S_n(p_1, \ldots, p_n)\}$$

Tropicalization in algebraic geometry

- starting with a polynomial f defining a hypersurface V in $(\mathbb{C}^*)^n$
- Maslov dequantization, given by a one-parameter family f_h with zero set V_h
- example: for $f(x) = \sum_k a_k x^k$, write $a_k = e^{b_k}$ and $x^k = e^{kt}$, and replace $v = \log(\sum_k e^{kt+b_k})$ with deformed $v_h = h \log(\sum_k e^{(kt+b_k)/h})$: this gives dequantized family $f_h(x) = \sum_k a_k^{1/h} x^k$
- *amoeba* obtained by mapping V_h to $\mathbb{R}^n$ under map $\mathrm{Log}_h(z_1, \ldots, z_n) = (h \log |z_1|, \ldots, h \log |z_n|)$
- limit $h \to 0$, subsets $\mathcal{A}_h \subset \mathbb{R}^n$ converge in Hausdorff metric to tropical variety $\mathrm{Tro}(V)$



Maslov dequantization and tropicalization of $f(x) = x^2 + ex + 1$
from `http://www.pdmi.ras.ru/~olegviro/dequant`



amoeba and tropicalization of $z_1 + z_2 + 1 = 0$ in $(\mathbb{C}^*)^2$

Maslov dequantization

- instead of deforming coefficients of polynomial see dequantization by deforming ring operations
- Maslov dequantization based on family of semirings $\mathbb{R}_+$

$$a \oplus_h b = h \log(e^{a/h} + e^{b/h}), \qquad a \odot b = a + b$$

  isomorphism to usual $\mathbb{R}_+$ through $x \mapsto x^h$ but in limit $h \to 0$ becomes the idempotent tropical semiring
- $\Delta \subset \mathbb{R}^n$ convex lattice polyhedron

$$\{\alpha_j\}_{j \in \Delta}, \quad \phi_h(x) = \oplus_h(\alpha_j + jx)$$

- Maslov dequantization can be expressed in terms of the operation $\oplus_{\mathrm{Sh},T}$, where the dequantization parameter h plays the role of the temperature T
- dequantization with respect to other entropy functionals?

Rényi entropy:

$$\mathrm{Ry}_\alpha(p_1, \ldots, p_n) := \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right)$$

$$\lim_{\alpha \to 1} \mathrm{Ry}_\alpha(p_1, \ldots, p_n) = \mathrm{Sh}(p_1, \ldots, p_n)$$

- lack of associativity of $x \oplus_S y$, when $S = \mathrm{Ry}_\alpha$

$$\mathrm{Ry}_\alpha(p) = \frac{1}{1-\alpha} \log(p^\alpha + (1-p)^\alpha)$$

measured by the transformation $(p_1, p_2, p_3) \mapsto (p_3, p_2, p_1)$

- in a commutative non-associative semiring K lack of associativity corrected by morphism

$$\begin{array}{ccccc} K \otimes K \otimes K & & \xrightarrow{A} & & K \otimes K \otimes K \\ \downarrow{\scriptstyle \oplus_w \otimes 1} & & & & \downarrow{\scriptstyle 1 \otimes \oplus_w} \\ K \otimes K & \xrightarrow{\oplus_w} & K & \xleftarrow{\oplus_w} & K \otimes K \end{array}$$

  which makes the diagram commutative
- morphism simply given by $A(x \otimes y \otimes z) = z \otimes y \otimes x$
- exactly the transformation $(p_1, p_2, p_3) \mapsto (p_3, p_2, p_1)$

$$\mathrm{Ry}_\alpha(p_1) + (1-p_1)\mathrm{Ry}_\alpha(\frac{p_2}{1-p_1}) =$$

$$\frac{1}{1-\alpha}\left(\log(p_1^\alpha + (1-p_1)^\alpha) + (1-p_1)\log\left((\frac{p_2}{1-p_1})^\alpha + (\frac{1-p_1-p_2}{1-p_1})^\alpha\right)\right) =$$

$$\frac{1}{1-\alpha}\log\left((p_1^\alpha + (1-p_1)^\alpha)\frac{(\frac{p_2}{1-p_1})^\alpha + (\frac{p_3}{1-p_1})^\alpha}{((\frac{p_2}{1-p_1})^\alpha + (\frac{p_3}{1-p_1})^\alpha)^{p_1}}\right) =$$

$$\frac{1}{1-\alpha}\log\left((\frac{p_1 p_2}{1-p_1})^\alpha + (\frac{p_1 p_3}{1-p_1})^\alpha + p_2^\alpha + p_3^\alpha\right) - \frac{p_1}{1-\alpha}\log\left((\frac{p_2}{1-p_1})^\alpha + (\frac{p_3}{1-p_1})^\alpha\right)$$

$$\frac{1}{1-\alpha}\log\left(\frac{(p_2^\alpha + p_3^\alpha)(p_1^\alpha + (1-p_1)^\alpha)}{(1-p_1)^\alpha}\right) - \frac{p_1}{1-\alpha}\log\left(\frac{(p_2^\alpha + p_3^\alpha)}{(1-p_1)^\alpha}\right)$$

$$= \frac{1}{1-\alpha}\left((1-p_1)\log(p_2^\alpha + p_3^\alpha) + \log(p_1^\alpha + (1-p_1)^\alpha) - \alpha(1-p_1)\log(1-p_1)\right)$$

On the other hand, we have

$$\mathrm{Ry}_\alpha(p_1+p_2)+(p_1+p_2)\mathrm{Ry}_\alpha(\frac{p_1}{p_1+p_2}) =$$

$$\mathrm{Ry}_\alpha(1-p_3)+(1-p_3)\mathrm{Ry}_\alpha(\frac{p_1}{1-p_3}) = \mathrm{Ry}_\alpha(p_3)+(1-p_3)\mathrm{Ry}_\alpha(\frac{p_1}{1-p_3}) =$$

$$\frac{1}{1-\alpha}\log((\frac{p_1p_3}{1-p_3})^\alpha+(\frac{p_2p_3}{1-p_3})^\alpha+p_1^\alpha+p_2^\alpha)-\frac{p_3}{1-\alpha}\log((\frac{p_1}{1-p_3})^\alpha+(\frac{p_2}{1-p_3})^\alpha)$$

$$=\frac{1}{1-\alpha}((1-p_3)\log(p_2^\alpha+p_1^\alpha)+\log(p_3^\alpha+(1-p_3)^\alpha)-\alpha(1-p_3)\log(1-p_3))$$

Non-extensive thermodynamics:

• gas of particles with chemical potentials $\log x$ and $\log y$ and Hamiltonian (p mole fraction)

$$\mathcal{H} = p \log x + (1-p) \log y$$

• partition function $Z = e^{-F_{\rm eq}}$ with $F_{\rm eq}$ equilibrium value of free energy at temperature $T = 1/\beta$

$$x \oplus_{\beta,S} y = \max_p (e^{TS(p)+p\log x+(1-p)\log y})$$

partition sum of a two state system with energies x and y

• Gibbs free energy $F = H - TS$ with S entropy and H enthalpy with $\min\{H - TS\}$ min of Gibbs free energy

• mixing can happen in non-Boltzmann thermodynamics (non-extensive) leading to non-associative thermodynamic semirings

• Extensive thermodynamics: independent subsystems A and B, combined system $A \star B$

$$S(A \star B) = S(A) + S(B)$$

• Non-extensive deformations (Tsallis)

$$S_q(A \star B) = S_q(A) + S_q(B) + (1-q) S_q(A) S_q(B)$$

Tsallis entropy:

$$\mathrm{Ts}_\alpha(p) = \frac{1}{\alpha - 1}(1 - p^\alpha - (1-p)^\alpha)$$

reproduces Shannon entropy $\alpha \to 1$

• Tsallis entropy uniquely determined by symmetry $S(p) = S(1-p)$, minima $S(0) = S(1) = 0$, and α-deformed extensivity

$$S(p_1) + (1-p_1)^\alpha S(\frac{p_2}{1-p_1}) = S(p_1+p_2) + (p_1+p_2)^\alpha S(\frac{p_1}{p_1+p_2})$$

Tsallis thermodynamic semiring

• Tsallis thermodynamic semiring: commutativity, unitarity and associativity of α-deformed $\oplus_{\beta,S,\alpha}$

$$x \oplus_{\beta,S,\alpha} y = \min_p \{ p^\alpha x + (1-p)^\alpha y - \frac{1}{\beta} \mathrm{Ts}_\alpha(p) \}$$

• for Tsallis entropy associativity of the thermodynamic semiring can be restored by a deformation of the operation $\oplus_{S,T}$ depending on deformation parameter α (also written as q for q-deformed)

• in previous physical interpretation this means replacing the energy functional

$$\mathcal{H} = \sum p_i E_i$$

with free energy of q-deformed thermodynamics

$$\mathcal{H}_q = \sum p_i^q E_i$$

• $\alpha \in \mathbb{R}$ and ϕ a continuous function such that $\phi(\alpha)(1-\alpha) > 0$ for $\alpha \neq 1$, with

$$\lim_{\alpha \to 1} \phi(\alpha) = 0$$

and such that $\exists\, 0 \leqslant a < 1 < b$ with ϕ differentiable on $(a, 1) \cup (1, b)$ and

$$\lim_{\alpha \to 1} \frac{d\phi(\alpha)}{d\alpha} < 0$$

• generalized Tsallis entropy:

$$\mathrm{Ts}_\alpha(p) = \frac{1}{\phi(\alpha)} (p^\alpha + (1-p)^\alpha - 1)$$

reproduces the Shannon entropy in the $\alpha \to 1$ limit

• (Suyari-Furuichi) generalized Tsallis entropy unique entropy functions that are commutative, have the L/R identity property, and satisfy the α-associativity condition

$$S(p_1) + (1-p_1)^\alpha S(\frac{p_2}{1-p_1}) = S(p_1+p_2) + (p_1+p_2)^\alpha S(\frac{p_1}{p_1+p_2})$$

• General idea: transform axiomatic characterizations of various entropy functionals into algebraic properties of corresponding thermodynamic deformations of min-plus algebras

Thermodynamic semirings of functions

- Ξ compact Hausdorff space
- $S = (S_\eta)$ family of information measures depending continuously on $\eta \in \Xi$
- $K = \mathbb{R}^{\min,+} \cup \{\infty\}$ and $C(\Xi, K)$ continuous functions with pointwise operations

$$x(\eta) \oplus_{T,S_\eta} y(\eta) = \min_{p\in[0,1]} (p\, x(\eta) + (1-p)\, y(\eta) - T\, S_\eta(p))$$

and ordinary pointwise sum as $\odot$

Kullback–Leibler divergence

- KL-divergence of two probability distributions

$$\mathrm{KL}(p|q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$$

  in binary form $P = (p, 1-p)$ and $Q = (q, 1-q)$

- smooth univariate binary statistical n-manifold $\mathcal{Q}$ is a set of binary probability distributions $\mathcal{Q} = (q(\eta))$ smoothly parametrized by $\eta \in \mathbb{R}^n$
- topological univariate binary statistical n-space $\mathcal{Q}$ is a set of binary probability distributions $\mathcal{Q} = (q(\eta))$ continuously parameterized by $\eta \in \Xi$, with Ξ a compact Hausdorff topological space
- first is setting of information geometry, second setting for multifractal dynamical systems
- semirings $\mathcal{R} = C^\infty(\mathcal{X}, K)$ or $\mathcal{R} = C(\mathcal{X}, K)$ with $\oplus_{T,KL(\cdot|q(\eta))}$ pointwise in η

multifractal systems

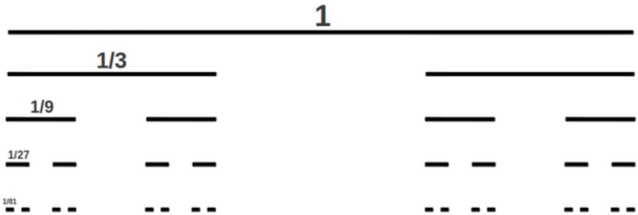
- Cantor set $\mathcal{X}$ identified with the one sided full shift space Σ_2^+ on the alphabet $\{0,1\}$

$$\eta = \eta_1\eta_2\eta_3\cdots\eta_n\cdots, \qquad \text{with } \eta_i \in \{0,1\}$$

  topologized with cylinder sets $\mathcal{X}(w) =$ sequences starting with the finite word w
- dynamical system with the shift map $\sigma(\eta) = \eta_2\eta_3\cdots$
- $a_n(\eta)$ denote the number of 1's that appear in the first n digits $\eta_1,\ldots,\eta_n$ of η
- when limit exists

$$q(\eta) = \lim_{n\to\infty} \frac{a_n(\eta)}{n}$$

- $\mathcal{Y} \subset \mathcal{X}$ set of points where limit exists



uniform middle-third Cantor set

- uniform Cantor set from contraction map f with contraction ratio λ
- with a Bernoulli measure μ_p for a given $0 < p < 1$

$$\mu_p(\mathcal{X}(w_1, \dots, w_n)) = p^{a_n(w)}(1-p)^{n-a_n(w)}$$

for cylinder sets

$$\mathcal{X}(w_1, \dots, w_n) = \{\eta \in \mathcal{X} \,|\, \eta_i = w_i,\ i = 1, \dots, n\}$$

- local dimension of $\mathcal{X}$ at a point $\eta \in \mathcal{Y}$ given by

$$d_{\mu_p}(\eta) = \frac{q(\eta)\log p + (1-q(\eta))\log(1-p)}{\log \lambda}$$

- local entropy of map f (shift σ) given by

$$h_{\mu_p,f}(\eta) = q(\eta)\log p + (1-q(\eta))\log(1-p)$$

- non-uniform Cantor set $\mathcal{X}$ with two contraction ratios λ_1 and λ_2 on the two intervals
- Lyapunov exponent of f is given by

$$\lambda_f(\eta) = q(\eta)\log\lambda_1 + (1-q(\eta))\log\lambda_2$$

- given Bernoulli measure μ_p on Cantor set $\mathcal{X}$ there is a set $\mathcal{Z} \subset \mathcal{X}$ of full measure $\mu_p(\mathcal{Z}) = 1$ for which $q(\eta) = p$
- uniform measure $\mu_{1/2}$: full measure subset $\mathcal{Z}_{1/2}$ with limit $q(\eta) = 1/2$ uniform distribution (fair coin case)
- stratify set $\mathcal{Y} \subset \mathcal{X}$ into level sets of $q(\eta)$: multifractal decomposition of Cantor set
- consider $C(\mathcal{Y}, K)$ with the $\oplus_{\mathrm{KL}_{q(\eta)}, T}$ with the Kullback–Leibler divergence $\mathrm{KL}(p; q(\eta))$

- for $\mathcal{Z} \subset \mathcal{Y}$ semiring $C(\mathcal{Z}, K)$ with $\oplus_{\mathrm{KL}_{q(\eta)}, T}$ is commutative iff $\mathcal{Z} \subset \mathcal{Z}_{1/2}$ is a "fair coin" subset
- involution that measures the lack of associativity and commutativity ($q \leftrightarrow 1 - q$)

$$(p_1, p_2, p_3; q) \mapsto (p_3, p_2, p_1; 1 - q)$$

- uniform case only associativity "up to a shift"

$$\mathrm{KL}(p_1; \frac{1}{2}) + (1 - p_1)\mathrm{KL}(\frac{p_2}{1 - p_1}; \frac{1}{2}) =$$

$$p_1 \log p_1 + p_2 \log p_2 + p_3 \log p_3 + \log 2 + (1 - p_1) \log 2$$

while

$$\mathrm{KL}(p_1 + p_2; \frac{1}{2}) + (p_1 + p_2)\mathrm{KL}(\frac{p_1}{p_1 + p_2}; \frac{1}{2}) =$$

$$p_1 \log p_1 + p_2 \log p_2 + p_3 \log p_3 + \log 2 + (1 - p_3) \log 2$$

- KL divergence to the uniform distribution = Shannon entropy up to a constant shift

KL divergence and marginal distributions

- p and q are two distributions, we denote by p_i and q_i their i-th marginal distribution, then $\mathrm{KL}(p|q) = \sum_i \mathrm{KL}(p_i|q_i)$

$$
\begin{aligned}
\mathrm{KL}(p|q) =& p_1 \cdots p_n \log \frac{p_1 \cdots p_n}{q_1 \cdots q_n} \\
&+ (1-p_1) p_2 \cdots p_n \log \frac{(1-p_1)p_2 \cdots p_n}{(1-q_1)q_2 \cdots q_n} \\
&+ \cdots + (1-p_1) \cdots (1-p_n) \log \frac{(1-p_1)\cdots(1-p_n)}{(1-q_1)\cdots(1-q_n)}
\end{aligned}
$$

$$
= p_1 \cdots p_n (\log \frac{p_1}{q_1} + \cdots + \log \frac{p_n}{q_n}) + \cdots + (1-p_1)\cdots(1-p_n)(\log \frac{1-p_1}{1-q_1} + \cdots + \log \frac{1-p_n}{1-q_n})
$$

$$
\begin{aligned}
=& p_1 \log \frac{p_1}{q_1} (p_2 \cdots p_n + (1-p_2)\cdots p_n + \cdots) \\
&+ (1-p_1) \log \frac{1-p_1}{1-q_1} (p_2 \cdots p_n + \cdots) + \cdots
\end{aligned}
$$

$$
= p_1 \log \frac{p_1}{q_1} ((1 + p_2 - p_2)(p_3 \cdots p_n + \cdots)) + \cdots
$$

$$
= p_1 \log \frac{p_1}{q_1} + (1-p_1) \log \frac{1-p_1}{1-q_1} + \cdots + (1-p_n) \log \frac{1-p_n}{1-q_n} = \sum_i \mathrm{KL}(p_i|q_i)
$$

- if sum of KL divergences of marginals is minimized total KL is also minimized

product of semirings and hyperfield structure

- semirings $\mathcal{R} = C(\{1, \dots, n\}, K) = K^{\otimes n}$
- want an n-ary probability distribution not n binary probability distributions
- want ordering on $\mathcal{R}$ that ensures trace is maximized

$$(x_1, ..., x_n) \to x_1 + ... + x_n \in K$$

- but such ordering does not uniquely determine a maximum between two tuples $\Rightarrow$ *non-well-defined addition on* K (multivalued): $(x_1, ..., x_n) + (y_1, ..., y_n)$ the set of tuples $(z_1, ..., z_n)$ with $z_i = x_i$ or y_i that maximize $z_1 + ... + z_n$ in the ordering on K
- this multivalued addition together with coordinate-wise multiplication defines a characteristic one hyperfield structure on $\mathcal{R}$
- information measures $S_1, ..., S_n$ over $K = \mathbb{R}^{\min,+} \cup \{\infty\}$
- for $x = (x_1, ..., x_n), y = (y_1, ..., y_n)$

$$x \oplus_{T,S_1,\dots,S_n} y = \min_{p_1,\dots,p_n} (p_1 x_1 + (1-p_1) y_1 - TS_1(p_1), ..., p_n x_n + (1-p_n) y_n - TS_n(p_n))$$

- the p_i as marginal probabilities, and the min operation is the multivalued hyperring addition
- if each S_i is the KL-divergence from some q_i, results of this operation are distributions with marginals $(p_1, ..., p_n)$ minimizing the KL-divergence to the marginals $(q_1, ..., q_n)$, subject to the *soft constraint* coming from the energy functional
$$\mathcal{H} = \sum p_i x_i + (1 - p_i) y_i$$
- lack of well-definedness of addition interpreted thermodynamically as non-uniqueness of equilibria, via existence of meta-equilibrium states
- when q_i uniform distribution addition is well-defined single-valued

Successor function

- encodes properties of thermodynamic semirings like measuring lack of associativity, commutativity
- $\lambda : K \times \mathbb{R} \to K$ is the Legendre transform of $TS : [0,1] \to \mathbb{R}$

$$\lambda(x, T) = x \oplus_S 0 \equiv \min_p (px - TS(p))$$

$$TS(p) = \min_x (px - \lambda(x, T))$$

- when S is concave/convex, we can recover it from λ hence from the semiring
- call λ successor function because 0 is the multiplicative identity and over general K we write this as $\lambda(x, T) = x \oplus_S 1$
- when multiplication distributes over addition

$$x \oplus_S y = \lambda(x - y, T) + y$$

- entropy function S has following properties:

1. commutativity $S(p) = S(1-p)$ (ie $\oplus_{S,T}$ commutative) iff

$$\lambda(x) - \lambda(-x) = x$$

2. left identity $S(0) = 0$ (ie $\oplus_S$ has left identity ∞) iff $\lambda(x) \leqslant 0$ and $\lim_{x\to\infty} \lambda(x) = 0$
3. right identity $S(1) = 0$ (ie $\oplus_S$ has right identity ∞) iff $\lambda(x) \leqslant x$ and $\lambda(x) \sim x$, as $x \to -\infty$
4. associativity iff

$$\lambda(x - \lambda(y)) + \lambda(y) = \lambda(\lambda(x-y) + y)$$

Successor for Shannon entropy

- in $\mathbb{R}^{\min,+} \cup \{\infty\}$

$$\lambda^{\rm Sh}(x, T) = -T \log(1 + e^{-x/T})$$

- in $\mathbb{R}^{\max,+}_{\geqslant 0}$

$$\lambda^{\rm Sh}(x, T) = (1 + x^{1/T})^T$$

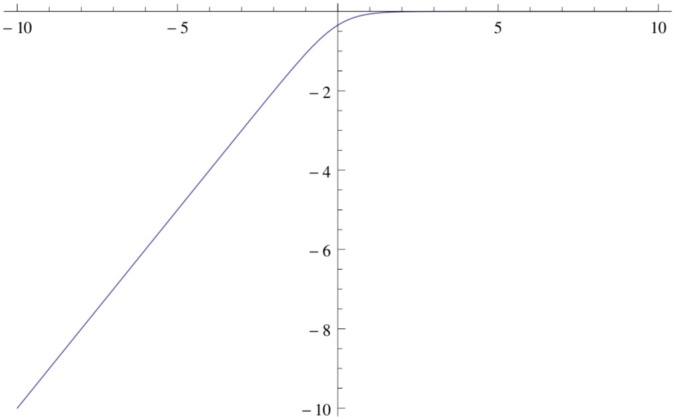
Successor for Kullback–Leibler divergence

- for $\mathbb{R}^{\min,+} \cup \{\infty\}$

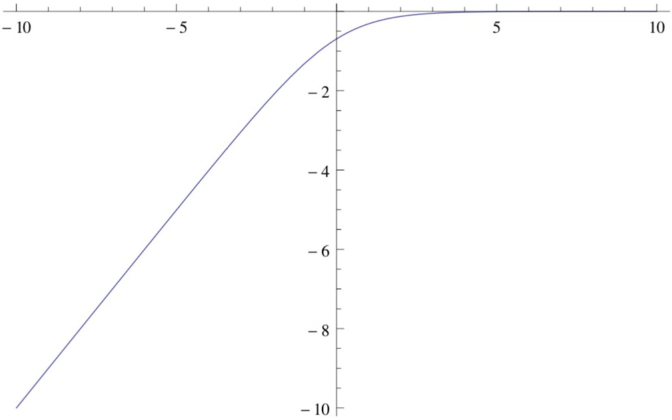
$$\lambda^{\rm KL}(x, T) = -T \log(1 + e^{-x/qT})$$

- for $\mathbb{R}^{\max,+}_{\geqslant 0}$

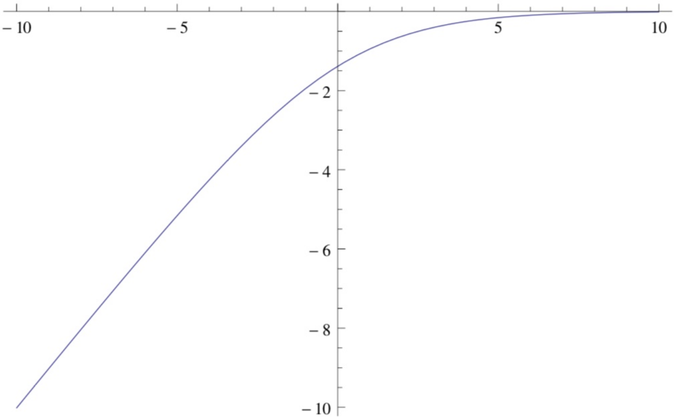
$$\lambda^{\rm KL}(x, T) = (1/(1-q)^{1/T} + (x/q)^{1/T})^T$$



successor function for the Shannon entropy with $T = 0.5$



successor function for the Shannon entropy with $T = 1$



successor function for the Shannon entropy with $T = 2$

Successor for Tsallis entropy

$$\lambda^{\mathrm{Ts}_\alpha}(x,T) = \begin{cases} 0 & |\frac{\alpha}{1-\alpha}| < x/T \\ g(x) & -|\frac{\alpha T}{1-\alpha}| < x/T < |\frac{\alpha}{1-\alpha}| \\ x & x/T < -|\frac{\alpha T}{1-\alpha}| \end{cases}$$

- $g(x)$ is given by applying Ts to the inverse of its derivative

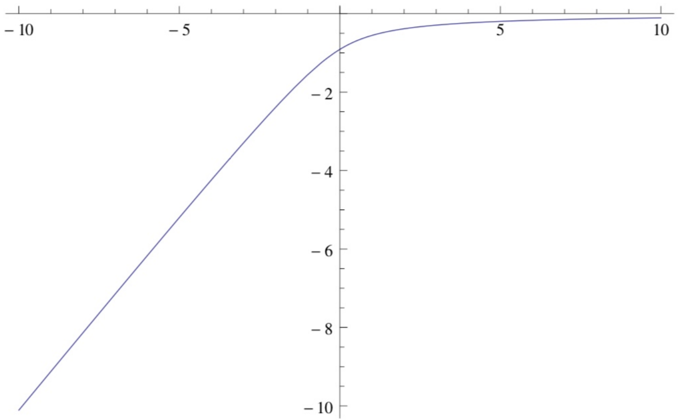
$$\frac{\partial \mathrm{Ts}}{\partial p} = \frac{\alpha}{1-\alpha}(p^{\alpha-1} - (1-p)^{\alpha-1})$$

that has range $[-|\frac{\alpha}{1-\alpha}|, |\frac{\alpha}{1-\alpha}|]$

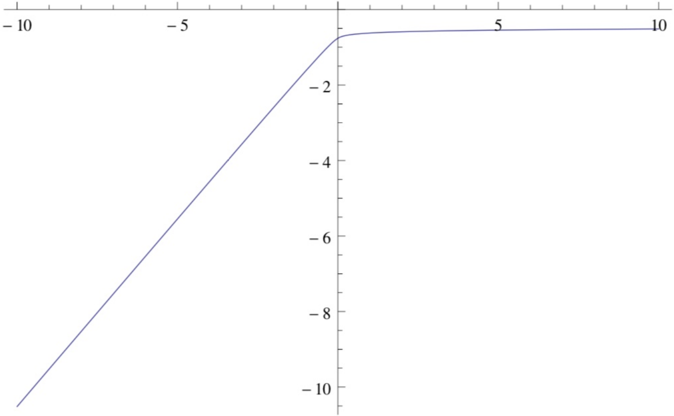
Successor for Rényi entropy

- applying Ry to inverse of its derivative (that now has range $\mathbb{R}$)

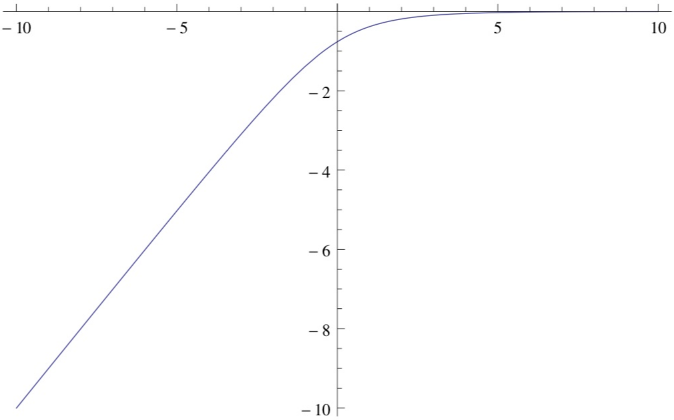
$$\frac{\partial \mathrm{Ry}}{\partial p} = \frac{\alpha}{1-\alpha}(p^{\alpha-1} + (1-p)^{\alpha-1})/(p^\alpha + (1-p)^\alpha).$$



successor function for the Tsallis entropy with $\alpha = 0.5$ and $T = 1$



successor function for the Rényi entropy with $\alpha = 0.1$ and $T = 1$



successor function for the Rényi entropy with $\alpha = 0.9$ and $T = 1$

Cumulant generating function

- random variable X
- $M_X(t)$ generating function of momenta of X

$$M_X(t) = \langle \exp(tX) \rangle = \sum_{m=0}^{\infty} \mu_m \frac{t^m}{m!}$$

- cumulants $\{\kappa_n\}$ of X coefficients of power series expansion of $\log M_X(t)$

$$\log M_X(t) = \sum_{n=0}^{\infty} \kappa_n \frac{t^n}{n!}$$

- information contained in cumulants or momenta is equivalent but cumulants are additive over independent variables

Cumulants and successor function

- $\lambda(x, T)$ successor function of a thermodynamic semiring K (assume it is analytic)
- function $-\lambda(x, T)/T$ is cumulant generating function of probability distribution for energy E in variable $-1/T = -\beta$

$$\kappa_n = \langle E^n \rangle_c$$

$$(-1)^{n+1} \frac{\partial^n}{\partial \beta^n} (\beta \lambda(x, T)) = \langle E^n \rangle_c$$

- partition function $Z(\beta) = \langle \exp(-\beta E) \rangle$
- Helmholtz free energy $F = -T \log \langle \exp(-E/T) \rangle$
- it is (up to $-1/T$ factor) cumulant generating function for random variable E
- Helmholtz free energy is Legendre transform of entropy so (up to $-1/T$ factor) the successor $\lambda(x, T)$

- for an arbitrary (concave and analytic) information measure

$$\lambda(x, T) - T\frac{\partial}{\partial T}\lambda(T, x) = \langle E \rangle = p_{\rm eq} x$$

with $p_{\rm eq} = p_T(x)$ equilibrium value of the mole fraction

- successor $\lambda(x, T) = \min_p(px - TS(p)) = p_T(x) - TS(p_T(x))$
- $p_T(x)$ satisfies

$$x/T = \frac{d}{dp}S(p_T(x))$$

$$\frac{\partial}{\partial T}\lambda(x/T) = x\frac{\partial}{\partial T}p(x/T) - S(p(x/T)) - T\frac{\partial}{\partial T}p(x/T)\frac{d}{dp}S(p(x/T))$$

which is just $-S(p(x/T))$

- so $p_T(x) = p(x/T)$ and similarly $\lambda(T, x) = \lambda(x/T)$
- explains effect of changing the temperature on $\oplus_{S,T}$
- get then $\frac{\partial}{\partial x}\lambda(x/T) = xp(x/T)$ hence well-known property of Legendre transform of smooth functions:

$$\lambda(x/T) = x\frac{\partial}{\partial x}\lambda(x/T) + T\frac{\partial}{\partial T}\lambda(x/T)$$

Operads

• Operad: objects $\mathcal{C}(j)$ in a symmetric monoidal category: parameter space of j-ary operations with composition maps

$$\gamma : \mathcal{C}(k) \otimes \mathcal{C}(j_1) \otimes \cdots \otimes \mathcal{C}(j_k) \to \mathcal{C}(j_1 + \cdots + j_k)$$

associative, unital, and equivariant under permutations

- unit: $e : I \to \mathcal{C}(1)$ (identity as unary operation) with identity compositions

$$\mathcal{C}(n) \simeq I \otimes \mathcal{C}(n) \stackrel{e\otimes 1}{\longrightarrow} \mathcal{C}(1) \otimes \mathcal{C}(n) \stackrel{\gamma}{\to} \mathcal{C}(n)$$

$$\mathcal{C}(n) \simeq \mathcal{C}(n) \otimes I^{\otimes n} \stackrel{1\otimes e^{\otimes n}}{\longrightarrow} \mathcal{C}(n) \otimes \mathcal{C}(1)^{\otimes n} \stackrel{\gamma}{\to} \mathcal{C}(n)$$

- associativity condition: composition γ of operations is associative (no ambiguity in how expressions involving composition operations are written without parentheses)
- Note: operations in $\mathcal{C}(j)$ are not necessarily associative
- *non-symmetric operad* if without the condition on the action of permutations

Algebras over an operad

way to realize operations in an operad $\mathcal{C}(j)$ as concrete operations: operations in $\mathcal{C}(j)$ have A-inputs (ie an input in $A^{\otimes j}$), output in A

- $\mathcal{C}$-algebra A: an object with Sym_j-equivariant maps

$$\mathcal{C}(j) \otimes A^{\otimes j} \to A,$$

thought of as actions, associative and unital

Note: operad structure closely related to trees and grafting trees, but there are different ways of organizing trees into an operad

Example: Operad of rooted trees

- operad $\mathcal{RT}$ with $\mathcal{RT}(n)$ the $\mathbb{Z}$-module generated by rooted trees T with n non-root vertices
- for an n-rooted tree T oriented towards the root, $\mathrm{In}(T,i) =$ set of incoming edges at the vertex i
- to define all the operad compositions sufficient to define for $1 \leq i \leq n$

$$\circ_i : \mathcal{RT}(n) \times \mathcal{RT}(m) \to \mathcal{RT}(n+m-1)$$

- operadic compositions then given by

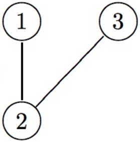
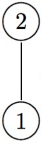
$$\gamma : \mathcal{RT}(n) \times \mathcal{RT}(k_1) \otimes \cdots \otimes \mathcal{RT}(k_n) \to \mathcal{RT}(k_1 + \cdots + k_n)$$

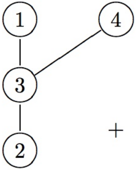
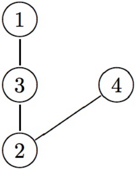
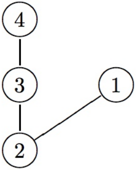
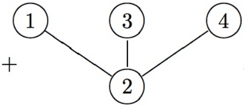
$$\gamma(T, S_1, \ldots, S_n) = (\cdots(T \circ_n S_n) \circ_{n-1} S_{n-1}) \cdots \circ_1 S_1)$$

- for $T \in \mathcal{RT}(n)$ and $S \in \mathcal{RT}(m)$

$$T \circ_i S = \sum_{f:\mathrm{In}(T,i)\to\{1,\ldots,m\}} T \circ_i^f S$$

composition along the vertex i of T: outgoing edge at i becomes outgoing edge at the root of S, incoming edges at i are grafted on vertices of S as specified by the map f

$T =$  and $S =$  .

$T \circ_2 S =$  $+$  $+$  $+$  .

$f\!:\!\begin{cases} 1 \to 2 \\ 3 \to 2 \end{cases}$ $\quad f\!:\!\begin{cases} 1 \to 2 \\ 3 \to 1 \end{cases}$ $\quad f\!:\!\begin{cases} 1 \to 1 \\ 3 \to 2 \end{cases}$ $\quad f\!:\!\begin{cases} 1 \to 1 \\ 3 \to 1 \end{cases}$

- Poincaré series of the operad $\mathcal{RT}$ of rooted trees

$$\Gamma_{\mathcal{RT}}(x) = \sum_{n\geq 1} \dim(\mathcal{RT}) \frac{(-x)^n}{n!}$$

- $\dim(\mathcal{RT}) = n^{n-1}$
- $\Gamma_{\mathcal{RT}}(x)$ is the inverse function of $x \mapsto -xe^{-x}$

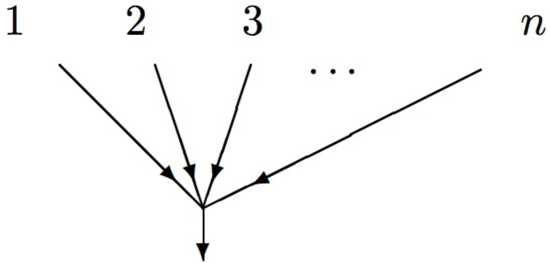
Example: A_∞-operad of planar rooted trees

- $A_\infty(n)$ is the linear span (over field $\mathbb{K}$ or over $\mathbb{Z}$) of planar rooted trees with n leaves (with labels $\{1,\ldots,n\}$)
- for $n = 1$ tree just a segment from root to single leaf
- symmetric group acts by relabeling the leaves
- operad composition

$$\gamma : A_\infty(n) \otimes A_\infty(k_1) \otimes \cdots \otimes A_\infty(k_n) \to A_\infty(k_1 + \cdots + k_n)$$

grafting n input leaves of $T \in A_\infty(n)$ to output roots of $T_i \in A_\infty(k_i)$

$\delta_n :=$



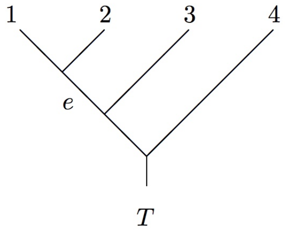
free operad generated by the corollas

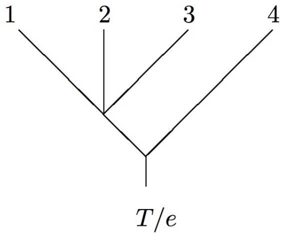
DG operad

- in grafting composition also introduce multiplication by a sign: grafiting root of T' to i-th leaf of T multiplied by $(-1)^{(\#E(T')-1)\cdot R_i(T)}$ with $R_i(T)$ = number of edges to the right or i-th leaf in T (strictly on the right of unique path from i-th leaf to root)
- reason for sign: the operad A_∞ also had a compatible differential (DG-operad)
- differential defined by edge contractions

$$dT := \sum_{T' \,:\, T = T'/e} \epsilon \, T'$$

$\epsilon = (-1)^{L(e)}$ with $L(e)$ = number of edges below and to the left of e





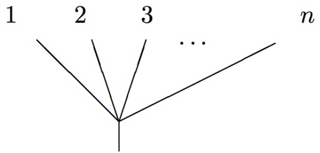
edge contraction operation on a tree

- differential satisfies $d^2 = 0$
- degree $\deg(T) = \#V(T) + 1 - n = \#E(T) + 1 - 2n$
- d has degree $+1$
- compatibility between operad structure and differential:

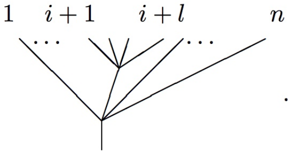
$$d(T' \circ_i T) = dT' \circ_i T + (-1)^{\deg(T')} T' \circ_i dT$$

- here still using the fact that the compositions $\circ_i$ determine all the operad compositions γ

d



$$= \sum_{\substack{k+l=n+1 \\ k,l\geq 2}} \sum_{i=0}^{k-1} (-1)^i$$



.

differential for corollas

- algebra over the A_∞ operad of planar rooted trees is a morphism of operads (compatible with all compositions)

$$\phi : A_\infty(n) \to \mathrm{End}_V(n)$$

where $\mathrm{End}_V(n) = \mathrm{Hom}(V^{\otimes n}, V)$ endomorphism operad

- compatibiity with differentials (morphisms of DG operads) if images $\phi(\delta_n)$ of corollas satisfy the differential as above
- same as requirement that $M_n(v_1, \ldots, v_n) = \phi(\delta_n)(v_1, \ldots, v_n)$ gives DG-algebra structure to V:

$$dM_n(v_1, \ldots, v_n) - (-1)^n \sum_{i=1}^{n} \epsilon(i) M_n(v_1, \ldots, dv_i, \ldots, v_n)$$

$$= \sum_{\substack{k+l=n+1 \\ k,l\geq 2}} \sum_{i=0}^{k-1} (-1)^{i+l(n-i-l)} \sigma(i) M_k(v_1, \ldots, v_i, M_l(v_{i+1}, \ldots, v_{i+l}), \ldots, v_n)$$

$\epsilon(i) = (-1)^{\sum_{k=1}^{i-1} \deg(v_i)}$ sign by moving d across $v_1, \ldots, v_{i-1}$; $\sigma(i) =$ sign by moving M_l across $v_1, \ldots, v_i$

Entropy Operads

- operad $\mathcal{P}$ of probabilities on finite sets $\mathcal{P}(j) = \Delta_j$ simplex
- operadic compositions

$$\gamma((p_i)_{i\in j}\otimes(q_{1l})_{l\in k_0}\otimes\cdots\otimes(q_{jl})_{l\in k_{j-1}}) = (p_i q_{il})_{l\in k_i, i\in j} \in \mathcal{P}(k_0+\cdots+k_{j-1})$$

- describe forming composites of subsystems

Algebra $\mathbb{R}_+$ over $\mathcal{P}$

• category $\mathbb{R}_{\geq 0}$ with a single object and morphisms $x \in \mathbb{R}_{\geq 0}$ with action of $\mathcal{P}$ trivial on unique object and on morphisms

$$(p_i)_{i\in j}\cdot(x_i)_{i\in j} = \sum_i p_i x_i$$

• J. Baez, T. Fritz, T. Leinster, *A characterization of entropy in terms of information loss*, Entropy 13 (2011) no. 11, 1945–1957.

Information Algebras (over the entropy operad)

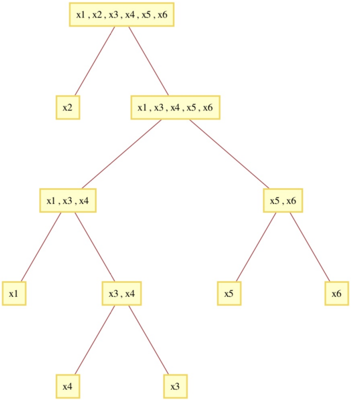
• object $\mathbb{R}_{\geq 0}$, morphisms $x \in \mathbb{R}_{\geq 0}$; action of operad $\mathcal{P}$: maps S from finite probabilities to non-negative real number with

1. For $p \in \mathcal{P}(n)$ and $q_i \in \mathcal{P}(m_i)$
$$S(p \circ (q_1, \dots, q_n)) = S(p) + \sum_i p_i S(q_i);$$
2. $S((1)) = 0$;
3. for $p \in \mathcal{P}(n)$ and $\sigma \in \mathrm{Sym}_n$
$$S(\sigma p) = S(p)$$
4. $S : \mathcal{P}(n) \to \mathbb{R}_{\geqslant 0}$ continuous

Characterizes entropy functionals (Khinchin axioms of Shannon entropy)

Binary guessing trees

- a general binary information measure, $S : [0,1] \to \mathbb{R}_{\geqslant 0}$ (not necessarily Shannon)
- assume S satisfies the L/R identity axioms
- build an information measure on ternary variables: X with values in $\{x_1, x_2, x_3\}$ guessed with binary questions in two ways
  1. Is $X = x_1$? If not, is $X = x_2$?
  2. Is $X = x_1$ or x_2? If yes, is $X = x_1$?
- counting possible permutations gives $2 \cdot 3! = 12$ possible ternary information measures
- S binary information measure with identity: for any $n \geqslant 2$, one-to-one correspondence between rooted binary trees with n leaves with labels in $\{1, \dots, n\}$ and n-ary information measures arising from S



Example: this tree corresponds to information measure $S_{\mathbf{T}}(p_1, \ldots, p_6) =$

$$S(p_2) + (1 - p_2)S(\frac{p_1 + p_4 + p_3}{1 - p_2}) + (p_1 + p_4 + p_3)S(\frac{p_1}{p_1 + p_4 + p_3})$$

$$+(p_4 + p_3)S(\frac{p_4}{p_4 + p_3}) + (p_5 + p_6)S(\frac{p_5}{p_5 + p_6})$$

- this shows how binary trees with n leaves and determine decision strategies and corresponding n-variable information measures
- conversely any decision strategy will consist of questions of the form is $X \in A$ for some subset $A \subset \{1, \dots, n\}$ and has to exhaust all possibilities so it will determine a binary rooted tree
- also an $(n, 2)$-tree $\mathbf{T}$ determines a canonical way of parenthesizing the expression $x_1 \oplus_S \cdots \oplus_S x_n$
- sufficient to consider case of $\mathbf{T}'$ where labels of leaves are $\{1, \dots, n\}$ from left to right (in planar embedding): for more general tree $\mathbf{T}$ there is a $\sigma \in S_n$ that gives the relabeling, then

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} := (x_{\sigma_{\mathbf{T}}(1)} \oplus_S \cdots \oplus_S x_{\sigma_{\mathbf{T}}(n)})_{\mathbf{T}'}$$

- then for the $\mathbf{T}$ with left-to-right labels construct inductively $(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}}$

- tree $\mathbf{T}_2$ with root vertex and two children leaves: $(x_1 \oplus_S x_2)_{\mathbf{T}_2} = x_1 \oplus_S x_2$
- if $\mathbf{T}$ has left-to-right labels, take L/R subtrees $\mathbf{L}$ and $\mathbf{D}$ at root: there is some $1 \leqslant r < n$ such that for $1 \leqslant j \leqslant r$ have $x_j \in L$, and for all $r < j < n$ have $x_j \in D$
- then inductively set

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = (x_1 \oplus_S \cdots \oplus_S x_r)_{\mathbf{L}} + (x_{r+1} \oplus_S \cdots \oplus_S x_n)_{\mathbf{D}}$$

Bracketing and multivariable information measures

- $(n, 2)$-tree $\mathbf{T}$ and binary information measure S with identity axiom
- relation between bracketing and the n-variable measure $S_{\mathbf{T}}$

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = \min_{\sum p_i = 1} (\sum p_i x_i - T S_{\mathbf{T}}(p_1, \ldots, p_n))$$

- first step: $\mathbf{T}$ labeled left-to-right labels and subtrees $\mathbf{L}$ with l leaves and $\mathbf{D}$ with d leaves at root:

$$S_{\mathbf{T}}(p_1, \ldots, p_l, p_{l+1}, \ldots, p_{l+d}) =$$

$$S(p_1 + \cdots + p_l) + (p_1 + \cdots + p_l) S_{\mathbf{L}}(\frac{p_1}{p_1 + \cdots + p_l}, \ldots, \frac{p_l}{p_1 + \cdots + p_l})$$

$$+ (p_{l+1} + \cdots + p_{l+d}) S_{\mathbf{D}}(\frac{p_{l+1}}{p_{l+1} + \cdots + p_{l+d}}, \ldots, \frac{p_{l+d}}{p_{l+1} + \cdots + p_{l+d}})$$

- second step: in this case also

$$(x_1 \oplus_S \cdots \oplus_S x_l \oplus_S x_{l+1} \oplus_S \cdots \oplus_S x_{l+d})_{\mathbf{T}}$$

$$= \min_p (p(x_1 \oplus_S \cdots \oplus_S x_l)_{\mathbf{L}} + (1-p)(x_{l+1} \oplus_S \cdots \oplus_S x_{l+d})_{\mathbf{D}} - TS(p))$$

- third step: induction on trees with less than n leaves

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = \min_p (p \min_{p_1+\cdots+p_l=1} (\sum p_i x_i - T S_{\mathbf{L}}(p_1,\ldots,p_l))$$

$$+(1-p) \min_{p_{l+1}+\cdots+p_{l+d}=1} (\sum p_i x_i - T S_{\mathbf{D}}(p_{l+1},\ldots,p_{l+d})) - T S(p))$$

- fourth step: change of variables $q_i = p p_i$, for each $i \in \{1,\ldots,l\}$, and $q_i = (1-p)p_i$, for each $i \in \{l+1,\ldots,l+d\}$
- with this have $q_1 + \cdots + q_l = p$ and $q_{l+1} + \cdots + q_{l+d} = 1-p$ so get

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = \min_{\sum q_i = 1} (\sum q_i x_i$$

$$-T((q_1+\cdots+q_l) S_{\mathbf{L}}(\frac{q_1}{q_1+\cdots+q_l},\ldots,\frac{q_l}{q_1+\cdots+q_l})$$

$$+(q_{l+1}+\cdots+q_{l+d}) S_{\mathbf{D}}(\frac{q_{l+1}}{q_{l+1}+\cdots+q_{l+d}},\ldots,\frac{q_{l+d}}{q_{l+1}+\cdots+q_{l+d}})$$

$$+S(q_1+\cdots+q_l)))$$

- by first step this is

$$\min_{\sum q_i=1}(\sum q_i x_i - TS_{\mathbf{T}}(q_1,\ldots,q_n)))$$

- for arbitrary labelings then use permutation σ and $p_i = q_{\sigma^{-1}(i)}$:

$$(x_{\sigma(1)} \oplus_S \cdots \oplus_S x_{\sigma(n)})_{\mathbf{T}} = \min_{\sum q_i=1}(\sum q_i x_{\sigma(i)} - TS_{\mathbf{T}}(q_1,\ldots q_n))$$

$$= \min_{\sum p_i=1}(\sum p_i x_i - TS_{\mathbf{T}}(p_{\sigma(1)},\ldots,p_{\sigma(n)}))$$

Example: same binary tree illustrated above

$$x_1 \oplus_S ((x_2 \oplus_S (x_3 \oplus_S x_4)) \oplus_S (x_5 \oplus_S x_6)) =$$

$$\min_{p_1}(p_1 x_1 + (1-p_1)((x_2 \oplus_S (x_3 \oplus_S x_4)) \oplus_S (x_5 \oplus_S x_6)) - TS(p_1))$$

$$= \min_{p_1}(p_1 x_1 + (1-p_1)\min_{p_2}(p_2(x_2 \oplus_S (x_3 \oplus_S x_4)) + (1-p_2)(x_5 \oplus_S x_6) - TS(p_2)) - TS(p_1))$$

$$= \min_{p_1,p_2}\Big(p_1 x_1 + (1-p_1)p_2 \min_{p_3}(p_3 x_2 + (1-p_3)(x_3 \oplus_S x_4) - TS(p_3))$$

$$+(1-p_1)(1-p_2)\min_{p_4}(p_4 x_5 + (1-p_4)x_6 - TS(p_4)) - T(S(p_1) + (1-p_1)S(p_2))\Big)$$

$$= \min_{p_1,p_2,p_3,p_4,p_5}(p_1 x_1 + (1-p_1)p_2 p_3 x_2 + (1-p_1)p_2(1-p_3)p_5 x_3$$

$$+(1-p_1)p_2(1-p_3)(1-p_5)x_4 + (1-p_1)(1-p_2)p_4 x_5 + (1-p_1)(1-p_2)(1-p_4)x_6$$

$$-T(S(p_1)+(1-p_1)S(p_2)+(1-p_1)p_2 S(p_3)+(1-p_1)(1-p_2)S(p_4)+(1-p_1)p_2(1-p_3)S(p_5)))$$

- then change of variables with $q_1 + \cdots + q_6 = 1$:

$$\begin{aligned}
p_1 &= q_1 \\
p_2 &= (q_2 + q_3 + q_4)/(1-q_1) \\
p_3 &= q_2/(q_2 + q_3 + q_4) \\
p_4 &= q_5/(q_5 + q_6) \\
p_5 &= q_3/(q_3 + q_4).
\end{aligned}$$

- change of variables gives

$$x_1 \oplus_S ((x_2 \oplus_S (x_3 \oplus_S x_4)) \oplus_S (x_5 \oplus_S x_6)) =$$

$$\min_{\sum q_i = 1} (\sum q_i x_i - T(S(q_1) + (1-q_1) S(\frac{q_2+q_3+q_4}{1-q_1})$$

$$+(q_2+q_3+q_4) S(\frac{q_2}{q_2+q_3+q_4}) + (q_3+q_4) S(\frac{q_3}{q_3+q_4}) + (q_5+q_6) S(\frac{q_5}{q_5+q_6}))$$

- then get by applying permutation $\sigma = (12)(34) \in S_6$

More general (non-binary) guessing trees

- for some $V \subset \mathbb{N}_{\geq 2}$ a family $\{S_n\}_{n\in V}$ of n-ary information measures
- satisfying coherence axiom: for $n > m$ if for all but $1 < i_1 < \cdots < i_m < n$, $p_j = 0$ then

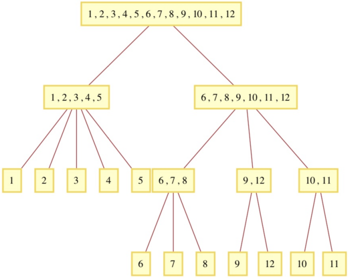
$$S_n(p_1, \ldots, p_n) = S_m(p_{i_1}, \ldots, p_{i_m})$$

- typical examples have the form

$$S_n(p_1, \ldots, p_n) = f(\sum_{1\leqslant i\leqslant n} g(p_i))$$

  for suitable functions f, g (all these satisfy coherence)
- for $v = \sup V$, can ask questions with up to v possible answers (instead of just binary)

- given $n, v \geqslant 2$, suppose for each $2 \leqslant j < v+1$ have j-ary information measure S_j (with coherence axiom)
- guessing strategies of n-ary random variables with questions with up to v possible answers
- in bijective correspondence with the set of (n, v)-trees: rooted trees with labelled leaves such that every vertex is either a leaf or has between 2 and v children



Entropy functionals and general planar rooted trees

- A collection $\mathcal{S} = \{S_n\}_{n\in\mathbb{N}}$ of n-ary entropy functionals S_n
- coherence condition:

$$S_n(p_1,\ldots,p_n) = S_m(p_{i_1},\ldots,p_{i_m}),$$

  whenever, for some $m < n$, we have $p_j = 0$ for all $j \notin \{i_1,\ldots,i_m\}$
- Shannon, Rényi, Tsallis entropies satisfy this condition
- collection $\mathcal{S} = \{S_n\}_{n\in\mathbb{N}}$ of coherent entropy functionals determines n-ary operations $C_{n,\beta,\mathcal{S}}$ on $\mathbb{R}\cup\{\infty\}$

$$C_{n,\beta,\mathcal{S}}(x_1,\ldots,x_n) = \min_p \{\sum_{i=1}^n p_i x_i - \frac{1}{\beta} S_n(p_1,\ldots,p_n)\}$$

  minimum taken over $p = (p_i)$, with $\sum_i p_i = 1$
- also write as $(x_1 \oplus_S \cdots \oplus_S x_n) := C_{n,\beta,\mathcal{S}}(x_1,\ldots,x_n)$
- in example tree above

$$(x_1 \oplus_S x_2 \oplus_S x_3 \oplus_S x_4 \oplus_S x_5) \oplus_S ((x_6 \oplus_S x_7 \oplus_S x_8) \oplus_S (x_9 \oplus_S x_{12}) \oplus_S (x_{10} \oplus_S x_{11}))$$

- more generally: n-ary operations $C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n)$ with $\mathcal{S}$ as above and $\mathbf{T}$ planar rooted trees with n leaves

$$C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n) = \min_p\{\sum_{i=1}^n p_i x_i - \frac{1}{\beta} S_{\mathbf{T}}(p_1,\ldots,p_n)\}$$

- with the $S_{\mathbf{T}}(p_1,\ldots,p_n)$ obtained from the S_j, for $j=2,\ldots,n$
- if root of (n,v)-tree $\mathbf{T}$ has sub-(l_j,v)-trees (resp. from left to right) $\mathbf{A}_1,\ldots,\mathbf{A}_m$, and the leaves of $\mathbf{T}$ are labeled left to right (and $L_j = l_1+\cdots+l_j$ and $L_0=0$)

$$S_{\mathbf{T}}(p_1,\ldots,p_n) =$$

$$\sum_{1\leqslant j\leqslant m} (p_{L_{j-1}+1}+\cdots+p_{L_j}) S_{\mathbf{A}_j}\left(\frac{p_{L_{j-1}+1}}{p_{L_{j-1}+1}+\cdots+p_{L_j}},\ldots,\frac{p_{L_j}}{p_{L_{j-1}+1}+\cdots+p_{L_j}}\right)$$

$$+S_m(p_{L_0+1}+\cdots+p_{L_1},\ldots,p_{L_{m-1}+1}+\cdots+p_{L_m})$$

- then also for $C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n) =: (x_1\oplus_S\cdots\oplus_S x_n)_{\mathbf{T}}$

$$(x_1\oplus_S\cdots\oplus_S x_n)_{\mathbf{T}} = \min_{\sum q_i=1}(q_1(x_1\oplus_S\cdots\oplus_S x_{l_1})$$

$$+\cdots+q_m(x_{l_1+\cdots+l_{m-1}+1}\oplus_S\cdots\oplus_S x_{l_1+\cdots+l_m})$$

$$-TS_m(q_1,\ldots,q_m)).$$

- same inductive argument then gives

$$C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n) = \min_{\sum p_i=1}(\sum p_i x_i - TS_{\mathbf{T}}(p_1,\ldots,p_n))$$

- indeed have

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = \min_{\sum q_i=1}(q_1 \min_{p_1+\cdots+p_{l_1}=1}(p_1x_1+\cdots+p_{l_1}x_{l_1} - TS_{\mathbf{A}_1}(p_1,\ldots,p_{l_1})+\cdots$$

$$+q_k \min_{p_{l_1+\cdots+l_{k-1}+1}+\cdots+p_{l_1+\cdots+l_k}=1}(\sum_{j=l_1+\cdots+l_{k-1}+1}^{l_1+\cdots+l_k} p_jx_j - TS_{\mathbf{A}_k}(p_{l_1+\cdots+l_{k-1}+1},\ldots,p_{l_1+\cdots+l_k}))$$

$$-TS_k(p_1+\cdots+p_{l_1},\ldots,p_{l_1+\cdots+l_{k-1}+1}+\cdots+p_n))$$

- for each $i \in \{1,\ldots,k\}$, and each $j \in \{l_1+\cdots+l_{i-1}+1,\ldots,l_1+\cdots+l_i\}$, with $l_0 = 0$, use substitution $\tilde{q}_j = q_i p_j$ with $\sum_{j=l_1+\cdots+l_{i-1}+}^{l_1+\cdots+l_i} \tilde{q}_j = q_i$:

$$(x_1 \oplus_S \cdots \oplus_S x_n)_{\mathbf{T}} = \min_{\sum \tilde{q}_j=1}(\sum \tilde{q}_j x_j - T((\tilde{q}_1+\cdots+\tilde{q}_{l_1})S_{\mathbf{A}_1}(\frac{\tilde{q}_1}{\tilde{q}_1+\cdots+\tilde{q}_{l_1}},\ldots)+\cdots$$

$$+(\tilde{q}_{l_1+\cdots+l_{k-1}+1}+\cdots+\tilde{q}_n)S_{\mathbf{A}_k}(\frac{\tilde{q}_{l_1+\cdots+l_{k-1}+1}}{\tilde{q}_{l_1+\cdots+l_{k-1}+1}+\cdots+\tilde{q}_n},\ldots))$$

- by first step (subtrees) this equals

$$\min_{\sum \tilde{q}_j = 1} (\sum \tilde{q}_j x_j - TS_{\mathbf{T}}(\tilde{q}_1, \ldots, \tilde{q}_n))$$

- then adjust order of leaves with a permutation, with $p_i = q_{\sigma^{-1}(i)}$:

$$(x_{\sigma(1)} \oplus_S \cdots \oplus_S x_{\sigma(n)})_{\mathbf{T}} = \min_{\sum q_i = 1} (\sum q_i x_{\sigma(i)} - TS_{\mathbf{T}}(q_1, \ldots q_n))$$

$$= \min_{\sum p_i = 1} (\sum p_i x_i - TS_{\mathbf{T}}(p_{\sigma(1)}, \ldots, p_{\sigma(n)}))$$

Entropy functionals and operad

- collection $\mathcal{S} = \{S_n\}_{n\in\mathbb{N}}$ of n-ary entropy functionals with coherence
- $\mathbb{R}$ as single-object topological category
- algebra over the A_∞ operad of rooted trees with operations

$$C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n) = \min_{\sum p_i=1}\left(\sum p_i x_i - T S_{\mathbf{T}}(p_1,\ldots,p_n)\right)$$

- satisfy additivity

$$C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_{j-1},x_j+y,x_{j+1},\ldots,x_n) =$$
$$C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_j,\ldots,x_n) + C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,y,\ldots,x_n)$$

- distributive property

$$y C_{n,\beta,\mathcal{S},\mathbf{T}}(x_1,\ldots,x_n) = C_{n,\beta,\mathcal{S},\mathbf{T}}(yx_1,\ldots,yx_n)$$

- also satifies (scaling of deformation variable)

$$(x_1 \oplus_S \cdots \oplus_S x_n)^{\alpha}_{\mathbf{T}}(T) = (x_1^{\alpha} \oplus_S \cdots \oplus_S x_n^{\alpha})_{\mathbf{T}}(\alpha T)$$

- additional relations between trees: $\mathbf{T}_1 \sim \mathbf{T}_2$ when $\forall x_i$

$$C_{n,\beta,\mathcal{S},\mathbf{T}_1}(x_1,\ldots,x_n) = C_{n,\beta,\mathcal{S},\mathbf{T}_2}(x_1,\ldots,x_n)$$

Information algebras and operad

- for $R = (\mathbb{R}_+, \max, +)$ and $\alpha_n : A_\infty(n) \to \mathbb{R}$
- in general, consider $A_\infty(n)$ vector space spanned by rooted trees with n leaves
- internal A_∞-algebra in R:
  1. for $\mathbf{T} \in A_\infty(n)$ and $\mathbf{A}_i \in A_\infty(k_i)$
     $$\alpha_{k_1+\cdots+k_n}\gamma(T, A_1, \ldots, A_n) = \alpha_n(\mathbf{T} + C_{n,\beta,\mathcal{S},\mathbf{T}}(\alpha_{k_1}(\mathbf{A}_1), \ldots, \alpha_{k_n}(\mathbf{A}_n))$$
  2. for all $\mathbf{T} \in A_\infty(n)$ and $\sigma \in \mathrm{Sym}_n$
     $$\alpha_n(\sigma\mathbf{T}) = \alpha_n(\mathbf{T})$$
  3. $\alpha_1|_{A_\infty(1)} = 0$
- values h_n of α_n on the (n, v)-tree with $n+1$ vertices

$$\alpha(\mathbf{T}) = h_n \oplus (\alpha(\mathbf{A}_1) \oplus_S \cdots \oplus_S \alpha(\mathbf{A}_n)) = \max(h_n, \alpha(\mathbf{A}_1) \oplus_S \cdots \oplus_S \alpha(\mathbf{A}_n))$$

- case of Shannon entropy: quotient of operad with exactly one class of (n, v)-trees for each n (by associativity and commutativity relations of Shannon)
- for Shannon entropy

$$\alpha(\mathbf{T}) = \max(h_n, (\alpha(\mathbf{A}_1)^{1/T} + \cdots + \alpha(\mathbf{A}_n)^{1/T})^T)$$

- so entropy functionals in classical probability can be understood as algebras over the A_∞-operad of rooted trees

Witt rings and thermodynamic semirings

- Witt ring $W(R)$ of a commutative ring R: as a group $(W(R), +_W)$ isomorphic to $(1 + tR[[t]], \times)$ and product $\star_W$ completely specified by

$$(1 - at)^{-1} \star_W (1 - bt)^{-1} = (1 - abt)^{-1}$$

  identities $1 = 1 + 0t + 0t^2 + \cdots$ for addition and $(1 - t)^{-1}$ for multiplication
- functoriality: ring homomorphisms $f : A \to B$ induce $W(f) : W(A) \to W(B)$
- built so that $W(\mathbb{F}_p)$ is the ring $\mathbb{Z}_p$ of p-adic integers
- elements $a \in \mathbb{Z}_p$ have usual p-adic expansion $a = a_0 + a_1 p + a_2 p^2 + \cdots$

- quotient map to residue field $\mathbb{Z}_p \to \mathbb{F}_p = \mathbb{Z}_p/p\mathbb{Z}_p$ has a unique *multiplicative* lift $\tau : \mathbb{F}_p \to \mathbb{Z}_p$ (map of multiplicative monoids) maps solution of $x^p = x$ in $\mathbb{F}_p$ to solution in $\mathbb{Z}_p$ congruent to x mod p

$$\pi \circ \tau = \mathrm{id}_K, \quad \tau(xy) = \tau(x)\tau(y) \quad \forall x, y \in \mathbb{F}_p$$

- can identify elements $a \in \mathbb{Z}_p$ uniquely through coordinates $a = \tau(a_0) + \tau(a_1)p + \tau(a_2)p^2 + \cdots$
- Witt vector $(\tau(a_0), \tau(a_1), \tau(a_2), \ldots)$
- these Teichmüller representatives are not $\{0, 1, \ldots, p-1\}$ but roots of $x^p - x = 0$
- Witt formula expresses sum of these representatives as

$$\tau(x) + \tau(y) = \tilde{\tau}(\sum_{\alpha \in I_p} w_p(\alpha, T) x^\alpha y^{1-\alpha})$$

$$I_p = \{\alpha \in \mathbb{Q} \cap [0,1] \,|\, p^n \alpha \in \mathbb{Z} \text{ for some } n\}$$

and $\tilde{\tau} : \mathbb{F}_p[[T]] \to \mathbb{Z}_p$ unique map with $\tilde{\tau}(xT^n) = \tau(x)p^n$ and $w_p(\alpha, T) \in \mathbb{F}_p[[T]]$

- the Witt formula has general form

$$x \oplus_w y = \sum_{s \in I_p} w_p(s) x^s y^{1-s}$$

- with coefficients

$$w_p(s) = \sum_{a/p^n = s} w(p^n, a) T^n$$

with $w(p^n, k) \in \mathbb{Z}/p\mathbb{Z}$, for $0 < k < p^n$, determined by addition formula

$$\tau(x) + \tau(y) = \tau(x+y) + \sum_{n=1}^{\infty} \tau \left(\sum w(p^n, k) x^{k/p^n} y^{1-k/p^n} \right) p^n$$

- as first observed by Connes-Consani the Shannon entropy semiring deformation gives an analog in characteristic one semirings

Entropy and thermodynamics in positive characteristics

- the universal sequence of the $w(p^n, k)$ can be seen as the characteristic p analog of the Shannon information
- non-extensive Tsallis entropy case also has a characteristic p analog in the form of q-deformations of the Witt ring
- see for instance:
  - Y.T. Oh, *q-Deformations of Witt–Burnside rings*, Math. Z. 257 (2007), N.1, 151–191.
  - M. Marcolli, Z. Ren, *q-deformations of statistical mechanical systems and motives over finite fields*, p-Adic Numbers Ultrametric Anal. Appl. 9 (2017), no. 3, 204–227

from classical to quantum: operadic structures

- from N. Combe, Yu.I. Manin, M. Marcolli, *Quantum operads*, in "Dialogues between physics and mathematics–C. N. Yang at 100", pp. 113–145, Springer, 2022.
- is there a way to extend the operad $\mathcal{P}$ of finite probabilities to quantum states?
- are there interesting algebras over operads in the quantum setting?
- what is the role of quantum information measures like von Neumann entropy?

density matrices and classical probabilities

given a density matrix ρ two natural classical probabilities associated to it:

1. $P = P(\rho)$ given by $p_i = \rho_{ii}$ diagonal elements
2. $\Lambda = \Lambda(\rho)$ eigenvalues λ_i of ρ (up to a choice of ordering: sort in non-increasing order)

properties:

- Schur lemma: sequence Λ of eigenvalues of hermitian matrix majorizes sequence P of diagonal entries, both sorted non-increasingly
- bistochastic matrix: since $\Lambda(\rho) \succ P(\rho)$ there is a bistochastic matrix B such that $P = B\Lambda$
- Shannon entropy non-decreasing $S(BP) \geq S(P)$
- Kullback–Leibler under bistochastic matrices $KL(BP|BQ) \leq K(P||Q)$
- uniform distribution $Q^{(N)} = (1/N)_{i=1}^N$ is a fixed point, $BQ^{(N)} = Q^{(N)}$
- Shannon entropy $KL(P|Q^{(N)}) = \log N - S(P)$

$\mathcal{Q}_P$-operad of quantum states

- $\mathcal{Q}_P(n) = \mathcal{M}^{(n)}$ convex set of density matrices
- composition laws

$$\gamma_P : \mathcal{Q}(n) \times \mathcal{Q}(k_1) \times \cdots \times \mathcal{Q}(k_n) \to \mathcal{Q}(k_1 + \cdots + k_n)$$

$$\gamma_P(\rho; \rho_1, \ldots, \rho_n) = \gamma(P(\rho); \rho_1, \ldots, \rho_n) = \begin{pmatrix} p_1\rho_1 & & & \\ & p_2\rho_2 & & \\ \vdots & & \cdots & \vdots \\ & & & p_n\rho_n \end{pmatrix}$$

- it is a non-unital, symmetric operad
- it restricts to unital operad $\mathcal{P}$ of classical probabilities on $\Delta_n \subset \mathcal{M}^{(n)}$

associativity of composition in $\mathcal{Q}_P$

- associativity condition is given by the identities

$$\gamma(\gamma(\rho^{(m)};\rho^{(n_1)},\ldots,\rho^{(n_m)});\rho^{(r_{1,1})},\ldots,\rho^{(r_{1,n_1})},\ldots,\rho^{(r_{m,1})},\ldots,\rho^{(r_{m,n_m})}) =$$

$$\gamma(\rho^{(m)};\gamma(\rho^{(n_1)};\rho^{(r_{1,1})},\ldots,\rho^{(r_{1,n_1})}),\ldots,\gamma(\rho^{(n_m)};\rho^{(r_{m,1})},\ldots,\rho^{(r_{m,n_m})}))$$

for $\rho^{(m)} \in \mathcal{Q}(m)$, $\rho^{(n_i)} \in \mathcal{Q}(n_i)$, $i = 1,\ldots,m$, and $\rho^{(r_{i,\ell_i})} \in \mathcal{Q}(r_{i,\ell_i})$ with $\ell_i = 1,\ldots,n_i$

- left-hand-side is

$$\gamma\left(\begin{pmatrix} \rho^{(m)}_{11}\rho^{n_1} & & & \\ & \rho^{(m)}_{22}\rho^{n_2} & & \\ \vdots & & \cdots & \vdots \\ & & & \rho^{(m)}_{mm}\rho^{n_m} \end{pmatrix}; \rho^{(r_{1,1})}, \ldots, \rho^{(r_{1,n_1})}, \ldots, \rho^{(r_{m,1})}, \ldots, \rho^{(r_{m,n_m})}\right) =$$

$$\begin{pmatrix} \rho^{(m)}_{11}\rho^{n_1}_{11}\rho^{(r_{1,1})} & & & & & & \\ & \ddots & & & & & \\ & & \rho^{(m)}_{11}\rho^{n_1}_{n_1 n_1}\rho^{(r_{1,n_1})} & & & & \\ & & & \ddots & & & \\ & & & & \rho^{(m)}_{mm}\rho^{n_1}_{11}\rho^{(r_{m,1})} & & \\ & & & & & \ddots & \\ & & & & & & \rho^{(m)}_{mm}\rho^{n_1}_{n_m n_m}\rho^{(r_{m,n_m})} \end{pmatrix}$$

which agrees with the right-hand-side

$$\gamma\left(\rho^{(m)}; \begin{pmatrix} \rho^{n_1}_{11}\rho^{(r_{1,1})} & & \\ & \ddots & \\ & & \rho^{n_1}_{n_1 n_1}\rho^{(r_{1,n_1})} \end{pmatrix}, \ldots, \begin{pmatrix} \rho^{n_m}_{11}\rho^{(r_{m,1})} & & \\ & \ddots & \\ & & \rho^{n_m}_{n_m n_m}\rho^{(r_{m,n_m})} \end{pmatrix}\right)$$

action of symmetric groups and equivariance

- permutations $\sigma_i \in \Sigma_{n_i}$ and $\sigma \in \Sigma_m$
- symmetric operad:
  1. first condition
  $$\gamma_P(\sigma(\rho); \rho_{\sigma^{-1}(1)}, \ldots, \rho_{\sigma^{-1}(m)}) = \tilde{\sigma}(\gamma_P(\rho; \rho_1, \ldots, \rho_m))$$
  where $\tilde{\sigma} \in \Sigma_{n_1+\cdots+n_m}$ permutation that splits into blocks of n_i indices and permutes blocks by σ
  2. second condition
  $$\gamma_P(\rho; \sigma_1(\rho_1), \ldots, \sigma_m(\rho_m)) = \hat{\sigma}(\gamma_P(\rho; \rho_1, \ldots, \rho_m))$$
  with $\hat{\sigma} \in \Sigma_{n_1+\cdots+n_m}$ permutation that acts on the i-th block of n_i indices as σ_i

- action of symmetric group Σ_n on $\mathcal{M}^{(n)}$ by $\sigma(\rho) = \sigma\rho\sigma^*$
- on diagonal matrices permutation of entries: agrees with Σ_n action on $\mathcal{P}(n)$
- induced action on Λ and P

$$\Lambda(\sigma\rho\sigma^*) = \Lambda(\rho) \quad \text{and} \quad P(\sigma\rho\sigma^*) = \sigma^* P(\rho)$$

- because $\sigma\rho\sigma^*$ and ρ same set of eigenvalues and both listed non-increasingly; for P have $p_i = \rho_{ii} = \mathrm{Tr}(\pi_i \rho)$ with 1-dim projection π_i and $\mathrm{Tr}(\pi_i \sigma\rho\sigma^*) = \mathrm{Tr}(\sigma^*\pi_i\sigma\rho) = \mathrm{Tr}(\pi_{\sigma^{-1}(i)}\rho)$
- so compatibility of operad $\mathcal{Q}_P$ with symmetric group action as

$$\gamma_P(\sigma(\rho); \rho_{\sigma^{-1}(1)}, \ldots, \rho_{\sigma^{-1}(m)}) = \gamma_P(\sigma^{-1}P(\rho); \rho_{\sigma^{-1}(1)}, \ldots, \rho_{\sigma^{-1}(m)})$$

is the same as $\tilde{\sigma}\gamma_P(\rho; \rho_1, \ldots, \rho_m))\tilde{\sigma}^*$ in $\mathcal{M}^{(n_1+\cdots+n_m)}$ and similar for second condition non-unital operad

- unit axiom is satisfied for $\rho = 1 \in \mathcal{Q}(1)$ with $\gamma_P(1; \rho) = \rho$
- it fails for $\rho_i = 1 \in \mathcal{Q}(1)$, where the composition gives instead $\gamma_P(\rho; 1, \ldots, 1) = P(\rho)$
- unit axiom is restored when restricting to $\Delta_n \subset \mathcal{M}^{(n)}$ classical probabilities where $\mathcal{Q}_P$ agrees with unital operad $\mathcal{P}$

$\mathcal{Q}_\Lambda$-operad of quantum states

- can do a similar construction of an operad with same $\mathcal{Q}_\Lambda(n) = \mathcal{M}^{(n)}$
- with composition given by

$$\gamma_\Lambda(\rho; \rho_1, \ldots, \rho_n) = \gamma(\Lambda(\rho); \rho_1, \ldots, \rho_n) = \begin{pmatrix} \lambda_1\rho_1 & & & \\ & \lambda_2\rho_2 & & \\ \vdots & & \cdots & \vdots \\ & & & \lambda_n\rho_n \end{pmatrix}$$

with λ_i the eigenvalues of ρ listed in non-increasing order

- the associativity condition for compositions holds as before because

$$\mathrm{Spec}\begin{pmatrix} \lambda_1\rho^{n_1} & & & \\ & \lambda_2\rho^{n_2} & & \\ \vdots & & \cdots & \vdots \\ & & & \lambda_m\rho^{n_m} \end{pmatrix} = \bigcup_i \lambda_i\,\mathrm{Spec}(\rho^{n_i})$$

- the unit axiom fails as before: $\gamma_\Lambda(1;\rho)=\rho$ holds but $\gamma_\Lambda(\rho;1,\ldots,1)=\Lambda(\rho)\neq\rho$
- the need to choose an ordering (non-increasing) of eigenvalues breaks the Σ_n-equivariance as $\Lambda(\sigma\rho\sigma^*)=\Lambda(\rho)$ so symmetric condition no longer satisfied
- $\mathcal{Q}_\Lambda$ is a non-unital non-symmetric operad
- when restricted to $\Delta_n\subset\mathcal{M}^{(n)}$ the operad $\mathcal{Q}_\Lambda$ only agrees with $\mathcal{P}$ up to permutations

different versions of non-unital operads

- for unital operads $\mathcal{O}$ can describe operad through compositions

$$\gamma : \mathcal{O}(n) \otimes \mathcal{O}(k_1) \otimes \cdots \otimes \mathcal{O}(k_n) \to \mathcal{O}(k_1 + \cdots + k_n)$$

- can also descibe $\mathcal{O}$ through insertion operations

$$\circ_i : \mathcal{O}(n) \otimes \mathcal{O}(m) \to \mathcal{O}(n+m-1)$$

- with constraints: for $1 \leq j \leq a$ and $b, c \geq 0$, with $X \in \mathcal{O}(a)$, $Y \in \mathcal{O}(b)$, and $Z \in \mathcal{O}(c)$

$$(X \circ_j Y) \circ_i Z = \begin{cases} (X \circ_i Z) \circ_{j+c-1} Y & 1 \leq i < j \\ X \circ_j (Y \circ_{i-j+1} Z) & j \leq i < b+j \\ (X \circ_{i-b+1} Z) \circ_j Y & j+b \leq i \leq a+b-1 \end{cases}$$

- then composition laws γ with associativity obtained by

$$\gamma(X, Y_1, \ldots, Y_n) = (\cdots (X \circ_n Y_n) \circ_{n-1} Y_{n-1}) \cdots \circ_1 Y_1)$$

- for non-unital operads these two descriptions no longer the same
- if described through the $\circ_i$ then also obtained via the γ operations
- but not always through that if described by the compositions γ these can be obtained from insertions $\circ_i$
- see M. Markl, *Operads and PROPs*, Handbook of algebra. Vol. 5, 87–140, Elsevier/North-Holland, 2008

- the operads $\mathcal{Q}_P$ and $\mathcal{Q}_\Lambda$ are also in the more restrictive class of non-unital operads
- the compositions γ are induced by insertions $\circ_i$
- for $\rho \in \mathcal{M}^{(n)}$ and $\rho' \in \mathcal{M}^{(m)}$ the insertion $\rho \circ_i \rho'$ in $\mathcal{M}^{(n+m-1)}$ is obtained by removing the i-th row and column of ρ and replacing them m rows and m columns, respectively, with
  - all the entries outside of the $m \times m$-block around the diagonal are zero
  - the $m \times m$-block around the diagonal is $p_i \rho'$
- then $(\cdots(\rho \circ_n \rho_n) \cdots \circ_1 \rho_1)$ produce exactly the matrix

$$\begin{pmatrix} p_1\rho_1 & & & \\ & p_2\rho_2 & & \\ \vdots & & \cdots & \vdots \\ & & & p_n\rho_n \end{pmatrix} = \gamma_P(\rho; \rho_1, \ldots, \rho_n)$$

- same thing for $\mathcal{Q}_\Lambda$ but with the $m \times m$-block around the diagonal given $\lambda_i \rho'$

range of the operad maps

- image of insertion map $\circ_i : \mathcal{Q}_P(n) \times \mathcal{Q}_P(m) \to \mathcal{Q}(n+m-1)$ consists of all density matrices $\rho \in \mathcal{M}^{(n+m-1)}$ that are block diagonal with one $(n-1) \times (n-1)$-block and one $m \times m$ block
- all block diagonal density matrices are in the image of some composition of insertion maps
- these are all the quantum states that decompose nontrivially into disjoint states with orthogonal ranges
- a block diagonal density matrix can be realized in several different ways through compositions of insertion maps

projective quantum measurements

- fixed fin dim Hilbert space $\mathcal{H}$ (choice of on basis)
- operators on density matrices are described by quantum channels
- special case: projective quantum measurement $\Pi = \{\Pi_i\}_{i=1}^n$ projectors $\Pi_i^* = \Pi_i = \Pi_i^2$ mutually orthogonal, $\Pi_i \Pi_j = \delta_{ij} \Pi_i$ and $\sum_i \Pi_i = 1$
- outcome of projective measurement Π on a quantum state $\rho \in \mathcal{M}^{(N)}$

$$\rho_i = \frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}(\Pi_i \rho)} \quad \text{with probability} \quad p_i = \mathrm{Tr}(\Pi_i \rho)$$

- quantum channel Π maps

$$\rho \mapsto \Pi(\rho) = \sum_i p_i \rho_i$$

- range in $\mathcal{M}^{(N)}$ of a composition of insertion maps specified by assigning $N = k_1 + \cdots + k_n$ and $\mathcal{M}_{k_1,\ldots,k_n} \subset \mathcal{M}^{(N)}$ block-diagonal density matrices with n blocks of size k_i
- projective measurement $\Pi = \{\Pi_i\}_{i=1}^n$, where Π_i is the orthogonal projection onto the span of the i-th subset of k_i basis elements
- quantum channel then maps $\Pi : \mathcal{M}^{(N)} \to \mathcal{M}_{k_1,\ldots,k_n}$, assigning to ρ the block-diagonal density matrix $\Pi(\rho) = \sum_i p_i \rho_i$
- different realizations as composition of insertion maps $\Leftrightarrow$ different realizations of this quantum channel through rooted trees

- τ planar rooted tree with n leaves decorated by integers $k_i \geq 1$, oriented from leaves to root
- at root vertex v_0 identity projector $\Pi^{(v_0)} = 1$
- v any vertex with set of incoming edges e at v, orthogonal projectors $\{\Pi^{(s(e))}\}_{t(e)=v}$ with $\sum_{e:t(e)=v} \Pi^{(s(e))} = \Pi^{(v)}$
- Π_i projectors at the leaves
- for $t(e) = v$

$$\rho^{(s(e))} = \frac{\Pi^{(s(e))}\rho^{(v)}\Pi^{(s(e))}}{\mathrm{Tr}(\Pi^{(s(e))}\rho^{(v)})},$$

with $\rho^{(v_0)} = \rho$

- quantum measurement Π_τ with outcomes ρ_i^τ with probabilities p_i^τ

$$p_i^\tau = \prod_w \mathrm{Tr}(\Pi_{i_w}^{(w)} \rho^{(w)})$$

product over the vertices on directed path from i-th leaf to root along with i_w idirection

- ρ_i^τ obtained by repeatedly computing $\rho^{(s(e))}$ from $\rho^{(t(e))}$ along the path from i-th leaf to root
- all the quantum channels Π_τ are the same quantum channel $\Pi : \mathcal{M}^{(N)} \to \mathcal{M}_{k_1,\dots,k_n}$

$$\prod_w \mathrm{Tr}(\Pi_{i_w}^{(w)} \rho^{(w)}) = \prod_\ell \frac{\mathrm{Tr}(\Pi^{(w_\ell)}\Pi^{(w_{\ell-1})}\cdots\Pi^{(w_0)}\rho)}{\mathrm{Tr}(\Pi^{(w_{\ell-1})}\cdots\Pi^{(w_0)}\rho)} = \prod_\ell \frac{\mathrm{Tr}(\Pi^{(w_\ell)}\rho)}{\mathrm{Tr}(\Pi^{(w_{\ell-1})}\rho)} = \mathrm{Tr}(\Pi_i\rho)$$

Π_i projection at leaf

$$\Pi_j^{(s(e))}\Pi_e^{(t(e))} = \Pi_j^{(s(e))}$$

as $\Pi_j^{(s(e))}$ projection onto a subspace of the range of $\Pi_e^{(t(e))}$

quantum entropy functionals

- family of quantum entropy functionals

$$S_n : \mathcal{M}^{(n)} \to \mathbb{R}$$

- consistency condition: S_n restricts to S_k, $k < n$, over $\mathcal{M}^{(k)} \subset \mathcal{M}^{(n)}$ ($n - k$ vanishing eigenvalues)
- Examples of consistent quantum entropy functionals:
  1. von Neumann entropy

$$\mathcal{N}(\rho) = -\mathrm{Tr}(\rho \log \rho)$$

  2. quantum Rényi entropy ($q \in \mathbb{R}_+^*$, $q \neq 1$)

$$Ry_q(\rho) = \frac{1}{1-q} \log \mathrm{Tr}(\rho^q)$$

  3. quantum Tsallis entropy ($q \in \mathbb{R}_+^*$, $q \neq 1$)

$$Ts_q(\rho) = \frac{1}{1-q}(\mathrm{Tr}(\rho^q) - 1)$$

Tree quantum entropy functionals

- tree τ with n leaves decorated by integers $k_i \geq 1$, $N = k_1 + \cdots + k_n$ together with coherent family $\{S_n\}$ of quantum entropies, determines entropy functional

$$S_\tau : \mathcal{M}^{(N)} \to \mathbb{R}$$

- if τ corolla with root vertex and n leaves

$$S_\tau(\rho) := S(P) + \sum_i p_i S(\rho_i)$$

  with $p_i = \mathrm{Tr}(\Pi_i \rho)$ and $\rho_i = \frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}(\Pi_i \rho)}$

- for von Neumann entropy by extensivity same as $\mathcal{N}(\Pi^\tau(\rho)) = \mathcal{N}(\sum_i p_i \rho_i)$
- Inductively assume S_τ constructed for all trees with less than n leaves
- subtrees τ_j, $j = 1, \ldots, m$, attached at root v_0, set L_j of leaves, $\# L_j < n$

- in quantum channel Π^τ have Π_j orthogonal projections with $\sum_j \Pi_j = 1$ along incoming edges e_j at root
- probabilities $p_j = \mathrm{Tr}(\Pi_j \rho)$ and density matrices $\rho_j = \frac{\Pi_j \rho \Pi_j}{\mathrm{Tr}(\Pi_j \rho)}$ at root vertices v_j of subtrees τ_j
- define

$$S_\tau(\rho) := S(P) + \sum_j p_j S_{\tau_j}(\rho_j)$$

  $S(P)$ Shannon entropy of classical probability $P = (p_j)$ and S_{τ_j} inductive entropy functionals of subtrees τ_j
- these completely specify S_τ
- for von Neumann by extensivity get $\mathcal{N}(\sum_i p_i \rho_i) = S(P) + \sum_i p_i \mathcal{N}(\rho_i)$ but not for other non-extensive entropies

Quantum channels

- quantum channel $\Phi : \mathcal{M}^{(N)} \to \mathcal{M}^{(N)}$ trace preserving completely positive map
- in Kraus form

$$\Phi(\rho) = \sum_i A_i \rho A_i^*$$

$\{A_i\}$ operators with $\sum_i A_i^* A_i = 1$ (but the $A_i^* A_i$ not necessarily projections)

- τ planar rooted tree with n leaves, oriented from leaves to root
- tree quantum channel C_A^τ assignment of operators $A = \{A_e\}_{e \in E(\tau)}$ edges of τ with at each vertex v condition

$$\sum_{e : t(e)=v} A_e^* A_e = 1$$

- channel C_A^τ acts on density matrices $\rho \in \mathcal{M}^{(N)}$ by

$$C_A^\tau(\rho) = \sum_{i=1}^n A_{\gamma_i} \rho A_{\gamma_i}^*$$

with $A_{\gamma_i} = A_{e_{i,1}} \cdots A_{e_{i,m_i}}$ along oriented path $\gamma_i e_{i,1}, \ldots, e_{i,m_i}$ from i-th leaf to root

- it is quantum channel

$$\sum_{i=1}^n A_{e_{i,m_i}}^* \cdots A_{e_{i,1}}^* A_{e_{i,1}} \cdots A_{e_{i,m_i}} = 1$$

as have inductively $\sum_{e\,:\,t(e)=v} A_e^* A_e = 1$ for paths of length one and reduce length by grouping sum by adjacent vertices

$$\sum_v \sum_{i\,:\,t(e_i)=v} A_{\gamma_v}^* A_{e_{i,1}}^* A_{e_{i,1}} A_{\gamma_v} = \sum_v A_{\gamma_v}^* A_{\gamma_v}$$

A_∞-operad $\mathcal{QC}$ of tree quantum channels

- $\mathcal{QC}(n) = span_{\mathbb{Z}}\{ C_A^\tau \,|\, \tau \in \mathcal{T}(n)\}$ with $\mathcal{T}(n)$ planar rooted trees
- operad composition laws

$$\gamma_{\mathcal{QC}} : \mathcal{QC}(n) \otimes \mathcal{QC}(k_1) \otimes \cdots \otimes \mathcal{QC}(k_n) \to \mathcal{QC}(k_1 + \cdots + k_n)$$

$$\gamma_{\mathcal{QC}}(C_A^\tau; C_{A_1}^{\tau_1}, \ldots, C_{A_n}^{\tau_n}) = C_{A\cup\{A_1,\ldots,A_n\}}^{\gamma_{\mathcal{T}}(\tau;\tau_1,\ldots,\tau_n)}$$

  with additional signs as in $\mathcal{T}$
- DG-structure (ϵ signs as in $\mathcal{T}$)

$$dC_A^\tau = \sum_{\tau' \,:\, \tau=\tau'/e} \epsilon \, C_{A'}^{\tau'}$$

- A' on τ' agrees with A on $\tau = \tau'/e$ for all edges non-adjacent to e, need to define A' on edges adjacent to e so that conditions at vertices still hold

- E_t set of edges of τ with target $t(e)$ in τ' and E_s the set of edges in τ with target $s(e)$ in τ'
- all same target vertex in τ so $\sum_{e' \in E_t \cup E_s} A^*_{e'} A_{e'} = 1$
- operators $B_t := \sum_{e' \in E_t} A^*_e A_e$ and $B_s := \sum_{e' \in E_s} A^*_e A_e$
- operators $\frac{1}{N_s} B_t$ and $A^*_{e'} A_{e'} + \frac{1}{N_s} B_t$ for $e' \in E_s$
- all these are positive, $\langle Bv, v \rangle \geq 0$ for all $v \in \mathcal{H}$
- so there are operators A and $\tilde{A}_{e'}$ with $B_s = A^* A$ and $A^*_{e'} A_{e'} + \frac{1}{N_s} B_t = \tilde{A}^*_{e'} \tilde{A}_{e'}$
- take A' to be $A'_e := A$ and $A'_{e'} := \tilde{A}_{e'}$ for $e' \in E_s$ and $A'_{e'} = A_{e'}$ for $e' \in E_t$

$$\sum_{e' : t(e') = s(e)} A'_{e'}{}^* A'_{e'} = \sum_{e' \in E_s} (A^*_{e'} A_{e'} + \frac{1}{N_s} B_t) = \sum_{e' \in E_s} A^*_{e'} A_{e'} + \sum_{e' \in E_t} A^*_{e'} A_{e'} = 1$$

$$\sum_{e' : t(e') = t(e)} A'_{e'}{}^* A'_{e'} = A'_e{}^* A'_e + \sum_{e' \in E_t} A^*_{e'} A_{e'} = \sum_{e' \in E_s} A^*_{e'} A_{e'} + \sum_{e' \in E_t} A^*_{e'} A_{e'} = 1$$

$\mathcal{QC}^+$-operad

- taking $\mathbb{Z}$-linear combinations of quantum channels allows for DG-structure but loses positivity
- if use convex combinations instead get operad but without differential
- $\mathcal{QC}^+(n) = \text{convex span}\{C_A^\tau \,|\, \tau \in \mathcal{T}(n)\}$
- compositions same as $\gamma_{\mathcal{QC}}$ (no signs)
- $\mathcal{M}^{(N)}$ algebra over the operad $\mathcal{QC}^+$
- operations $\alpha : \mathcal{QC}^+(n) \otimes \mathcal{M}^{(N)\otimes n} \to \mathcal{M}^{(N)}$

$$\alpha(C_A^\tau; \rho_1, \ldots, \rho_n) = \sum_{i=1}^n p_i \tilde{\rho}_i$$

$$\tilde{\rho}_i = \frac{A_{\gamma_i} \rho_i A_{\gamma_i}^*}{\mathrm{Tr}(A_{\gamma_i}^* A_{\gamma_i} \rho_i)} \quad \text{and} \quad p_i = \mathrm{Tr}(A_{\gamma_i}^* A_{\gamma_i}),$$

- $A_{\gamma_i} = A_{e_{i,1}} \cdots A_{e_{i,m_i}}$, along path $\gamma_i = e_{i,1}, \ldots, e_{i,m_i}$ from i-th leaf to root