Realizability and Extension of Measures: Classical and Quantum

Joel Lebowitz
Rutgers University

November 2013

I. Truncated Moment Problem : One Dimension
(with Tobias Kuna, Maria Infusino and Eugene Speer)

II. Truncated Moment Problem: Infinite Dimensions
(with Emanuele Caglioti, Ovidiu Costin, Tobias Kuna and Eugene Speer)

III. Extension of Measures: Classical
(with Shelley Goldstein, Tobias Kuna and Eugene Speer)

IV. Extension of Density Matrices: Quantum
(with Eric Carlen and Elliott Lieb)

I. Truncated moment problem : One random variable $x \in K$

I start with the following question :
given an n-tuple of real numbers $(m_1, m_2 \cdots, m_n)$,
does there exist a probability measure
$\mu(dt)$ on $K \subset \mathbb{R}$ such that $m_0 = 1$ and

$$\int_K t^k \mu(dt) = m_k, \quad k = 1, 2 \cdots n \tag{1}$$

Examples of necessary and sufficient conditions for (1) to hold with m_1, m_2, m_3 given are

1. $K = \mathbb{R}_+ = [0, \infty)$:
$m_1 \geq 0, m_2 - m_1^2 \geq 0, \frac{m_3}{m_1} - (\frac{m_2}{m_1})^2 \geq 0$, and
$m_2 = m_1^2 \Rightarrow m_k = m_1^k$

2. $K = \mathbb{N} = 0, 1, 2, \cdots$
$m_1 \geq 0, \quad m_2 - m_1^2 \geq \theta_1(1 - \theta_1)$

$\frac{m_3}{m_1} - (\frac{m_2}{m_1})^2 \geq \theta_2(1 - \theta_2)$

where $\theta_1, \theta_2 \in [0, 1)$ are the fractional parts of m_1 and m_2/m_1, respectively.

Case 1, is part of the standard truncated moment problem considered by many. There are known conditions for realizability (consistency) in terms of positivity of some Hankel matrices for all n.

Case 2, which corresponds, for example, to the statistics of the number of particles (or people) in a given region, is less well studied and much more complicated. There is no explicit formula for necessary and sufficient conditions when $n \geq 4$.

For $n = 2$ the condition $m_2 - m_1^2 \geq \theta_1(1 - \theta_1)$ is due to Percus and Yamada and goes, in the classical fluids literature under the name of Yamada condition. The proof is straightforward.

To minimize the variance $m_2 - m_1^2$, when $m_1 = k + \theta_1$ the measure has to be concentrated on k and $k+1$ with probabilities $(1 - \theta_1)$ and θ_1 respectively.
This gives $m_2 - m_1^2 = \theta_1(1 - \theta_1)$.

For $n = 3$, the situation is a bit more complicated but in the end the minimal values are obtained
if the measure is concentrated on three points $0, k_1, k_2$.

For $n = 4$, all we have explicitly are lower and upper bounds; m_4 has to be greater than it is for $\mathbb{R}_+$ where, the positivity of the Hankel matrix

$$\begin{pmatrix} m_0 & m_1 & m_2 \\ m_1 & m_2 & m_3 \\ m_2 & m_3 & m_4 \end{pmatrix}$$

gives

$$m_4 - m_2^2 \geq \frac{(m_3 - m_2 m_1)^2}{(m_2 - m_1^2)}$$

We also have an upper bound μ_4, such that $m_4 > \mu_4(m_1, m_2, m_3)$ is sufficient for realizability but no explicit necessary and sufficient condition. One can, in principle compute such a condition for each $m_1, \ldots, m_n$, but the complexity grows exponentially (?) in n.

More generally; let $\mathbf{t} = (t_1, \ldots, t_N) \in K \subset \mathbb{R}^N$. Then given

$$\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$$

where $\mathbf{m}_l = \{m_{l_1, l_2, \ldots, l_N}\}$, $l_j \geq 0$, $\sum_{j=1}^{N} l_j = l$, we ask, does there exist $\mu(dt_1, \ldots, dt_N)$ such that

$$\mathbf{m}_l = \iint_K t_1^{l_1} \ldots t_N^{l_N} \mu(\mathbf{dt})$$

We know very little about this for $N > 1$, so let us go on to the case $N = \infty$.

II. Point Processes with Specified Low Order Correlations

This is a generalization of the truncated moment problem, which arises naturally in statistical mechanics. In fact this was the origin of our interest in these problems.

Let $\eta(\mathbf{r})$, $\mathbf{r} \in \Omega$, be a random empirical field describing a point process in a domain $\Omega \subset \mathbb{Z}^d$ or $\mathbb{R}^d$. Then

$$\eta(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{x}_i),$$

where the $\mathbf{x}_i$ are the positions of the points of the process, with $\mathbf{x}_i \neq \mathbf{x}_j$ for $i \neq j$, distributed according to some measure μ defined on the family of all locally finite collections of points in Ω.

Here δ is either the Dirac or the Kronecker delta function, depending on whether we are in the continuum or on the lattice; in the latter case $\eta(\mathbf{r})$ has value 0 or 1.

Depending on context the $\mathbf{x}_i$ can represent the positions of particles in a fluid or of the stars in the sky, the occurrence times of members of a train of neural spikes, or more generally the space-time locations of the events of some specified physical process.

The correlation functions $\rho_k(\mathbf{r}_1, ..., \mathbf{r}_k)$ are defined via averages, with respect to μ, of products of $\eta(\mathbf{r})$'s involving distinct particles:

$$\begin{aligned} \rho_1(\mathbf{r}_1) &= \langle \eta(\mathbf{r}_1) \rangle, & (2) \\ \rho_2(\mathbf{r}_1, \mathbf{r}_2) &= \langle \eta(\mathbf{r}_1)\eta(\mathbf{r}_2) \rangle - \rho_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2), & (3) \end{aligned}$$

and in general

$$\rho_k(\mathbf{r}_1, \ldots, \mathbf{r}_k) = \left\langle \sum_{i_1 \neq i_2 \neq \cdots \neq i_k} \prod_{j=1}^{k} \delta(\mathbf{r}_j - \mathbf{x}_{i_j}) \right\rangle. \qquad (4)$$

Note that on the lattice $\rho_k(\mathbf{r}_1, ..., \mathbf{r}_k)$ vanishes when $\mathbf{r}_i = \mathbf{r}_j$.

For translation invariant processes, we shall write
$\rho_1(\mathbf{r}_1) = \rho$, $\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 g(\mathbf{r}_2 - \mathbf{r}_1)$,
$\rho_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \rho^3 g_3(\mathbf{r}_2 - \mathbf{r}_1, \mathbf{r_3} - \mathbf{r}_1)$, etc.
The function $g(\mathbf{r})$ is known in the fluids literature, where it is additionally assumed that g is a function only of $|\mathbf{r}|$, as the *radial distribution function*. We shall also assume generally that $\rho_k(\mathbf{r}_1, \ldots, \mathbf{r}_k) \to \rho^k$ when $|\mathbf{r}_i - \mathbf{r}_j| \to \infty$ for all i, j with $1 \leq i < j \leq k$.

We study the following infinite dimensional truncated moment problem.
Suppose we are given functions, say
$f_1(\mathbf{r}_1), f_2(\mathbf{r}_1, \mathbf{r}_2) \ldots f_n(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ does there exist an underlying point process with correlations $\rho_j = f_j, \quad j = 1, \ldots n$, and, if so, what can we say about it?

The given putative correlations f_j may come from averaging and smoothing of observations, as in the study of neural spike trains, or from some approximate theory, such as the Percus-Yevick equation for the radial distribution function of a classical fluid. They may also just express target correlations for a material or process with certain desired properties .

We observe that if the correlations $\rho_j = f_j$ can be realized for some density ρ, then they can also be realized, for the same functions $g_2, \ldots, g_k$, for any ρ', $0 \leq \rho' < \rho$. To see this, the new measure is constructed by independently choosing to delete or retain each point in a configuration, keeping a point with probability ρ'/ρ.

In this light it is thus natural to pose the realizability problem in the following form: given the g_j, $j = 2, \ldots, k$, what is the least upper bound $\bar{\rho}$ of the densities for which they can be realized? It is of course possible in the continuum case to have $\bar{\rho} = \infty$; for example, if $g_j = 1$ for $j \leq 2 \leq k$ then for any density $\rho > 0$ a Poisson process realizes the correlations. For the lattice systems considered here, on the other hand, we always have $\bar{\rho} \leq 1$.

Lacking a full answer to this question, one may of course ask rather for upper and lower bounds on $\bar{\rho}$. A lower bound $\bar{\rho} \geq \rho_0$ may be obtained as we will show by the construction of a process at a density $\rho_0 > 0$. Upper bounds on $\bar{\rho}$ may be obtained from necessary conditions for realizability, some of which are described below.

Necessary Conditions

Clearly, from (4), realizability requires that

$$\rho_j(\mathbf{r}_1, \ldots, \mathbf{r}_j) \geq 0, \qquad j = 1, \ldots, k. \tag{5}$$

We also know that the covariance matrix of the field $\eta(\mathbf{r})$,

$$\begin{aligned} S(\mathbf{r}_1, \mathbf{r}_2) &= \langle \eta(\mathbf{r}_1)\eta(\mathbf{r}_2)\rangle - \langle \eta(\mathbf{r}_1\rangle\langle \eta(\mathbf{r}_2)\rangle \\ &= \rho_2(\mathbf{r}_1, \mathbf{r}_2) + \rho_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2) - \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2) \end{aligned} \tag{6}$$

must be positive semi-definite, which implies that:

$$\int_\Lambda \rho_1(\mathbf{r}_1)\, d\mathbf{r}_1 + \int_\Lambda \int_\Lambda e^{i\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \left[\rho_2(\mathbf{r}_1, \mathbf{r}_2) - \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\right] d\mathbf{r}_1 d\mathbf{r}_2 \geq 0; \tag{7}$$

for any $\Lambda \subset \mathbb{R}^d(\mathbb{Z}^d)$

In the translation invariant case we must have the non-negativity of the infinite volume structure function $\widehat{S}(\mathbf{k})$:

$$\widehat{S}(\mathbf{k}) \equiv \rho + \rho^2 \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{r}} \, [g(\mathbf{r}) - 1] \, d\mathbf{r} \geq 0. \tag{8}$$

Here I assume $\int_{\mathbb{R}^d} |g(\mathbf{r}) - 1| \, d\mathbf{r} < \infty$; otherwise (8) holds in the sense of generalized functions. There are corresponding conditions on the torus $\mathbb{T}^d$, the lattice $\mathbb{Z}^d$, and the periodic lattice.

There are also necessary conditions corresponding to case 2 above: if N_Λ denotes the number of particles in a region $\Lambda \subset \Omega$, then the variance V_Λ of N_Λ,

$$\begin{aligned} V_\Lambda &\equiv \int_\Lambda \int_\Lambda S(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \int_\Lambda \rho_1(\mathbf{r}_1)\, d\mathbf{r}_1 + \int_\Lambda \int_\Lambda [\rho_2(\mathbf{r}_1, \mathbf{r}_2) - \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)]\, d\mathbf{r}_1 d\mathbf{r}_2, \quad (9) \end{aligned}$$

must satisfy the previous case (2) Yamada condition

$$V_\Lambda \geq \theta(1 - \theta), \quad (10)$$

The above conditions follow from the more general uncountable number of necessary conditions. In summary these say that, given any functions $f_2(\mathbf{r}_1, \mathbf{r}_2)$, $f_1(\mathbf{r})$ and constant f_0 such that, for any n points $\mathbf{r}_1, \ldots, \mathbf{r}_n$ in Λ,

$$\sum_{i \neq j} f_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i f_1(\mathbf{r}_i) + f_0 \geq 0,$$

then we must have

$$\iint\limits_{\Lambda\times\Lambda} \rho_2(\mathbf{r}_1, \mathbf{r}_2) f_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 + \int_\Lambda \rho_1(\mathbf{r}) f_1(\mathbf{r}) d\mathbf{r} + f_0 \geq 0, \tag{11}$$

for all $\Lambda \subset \mathbb{R}^d$.

We prove in fact that in the case $k = 2$, i.e. for the case that only ρ_1 and ρ_2 are given, (11) is also a sufficient condition for realizability under some assumptions on the point process, e.g, existence of a hard core.

This has been extended recently by Raphaël Lachièze-Rey and Ilya Molchanov to a weaker condition on $g(r) \to 0$ as $r \to 0$.

When there are no restrictions on the number of particles in a given region sufficiency only holds under some additional restriction. There are similar results for $k > 2$.

Note that in the case $k = 2$ all restrictions on ρ and g beyond those arising from nonnegativity of ρ and of the covariance matrix S of (6) are due to the fact that we want $\eta(\mathbf{r})$ to be a point process, since we can always find a Gaussian process realizing any ρ_1, ρ_2 with $S > 0$.

We also note that for $g(\mathbf{r}) \leq 1$ one has

$$\widehat{S}(\mathbf{k}) \geq \widehat{S}(\mathbf{0}) = \lim_{\Lambda \to \infty} \frac{1}{|\Lambda|} V_{\Lambda}. \tag{12}$$

In general $\widehat{S}(0) = 0$ implies that the variance V_{Λ} is growing slower than the volume. Processes with this property are called *superhomogeneous* and are of independent interest.

To prove the realizability of a given translation invariant ρ and $g(\mathbf{r})$, $\mathbf{r} \in \mathbb{R}^d$, for sufficiently small ρ we extend results given by Ambartzumian and Sukiasian. Given ρ and $g(\mathbf{r})$ they generated higher order correlation functions ρ_n, $n = 1, 2, 3, \ldots$. via the following ansatz.

$$\rho_n(\mathbf{r}_1, \ldots, \mathbf{r}_n) \;=\; \rho^n \prod_{1 \le i < j \le n} g(\mathbf{r}_i - \mathbf{r}_j), \qquad (13)$$

This solves the realization problem for ρ and g iff one can show that these ρ_n actually realize a point process in $\mathbb{R}^d(\mathbb{Z}^d)$. To do this requires some manipulations involving Ruelle's and Penrose's conditions for convergence of the fugacity expansion for a system with pair potential $\phi(r) = -\log g(r)$.

The end result is that this construction works whenever

$$\rho \leq (e^{\Phi+1} B)^{-1} \tag{14}$$

where

$$B = \int_{\mathbb{R}^d} |g(r) - 1| d\mathbf{r}$$

and Φ is the Ruelle stability condition

$$\sum_{i=1}^{N} (-\log g(\mathbf{x}_0 - \mathbf{x}_i)) \geq -N\Phi \tag{15}$$

whenever

$$g(\mathbf{x}_i - \mathbf{x}_j) > 0, \quad i, j = 0, \ldots, N \tag{16}$$

Note that if $g(\mathbf{r}) \leq 1$ then (15) is automatically satisfied with $\Phi = 0$. This was the case considered by Ambartzumian and Sukiasian.

Many will recognize (14) as the lower-bound on the radius of convergence of the Mayer fugacity expansion, which is indeed where it comes from.

The result also holds for $\mathbb{Z}^d$. It has also been extended by us to the case where one also specifies g_3 and when ρ_1, ρ_2, ρ_3 are not translation invariant. The value of $\overline{\rho}$ is much smaller then. The existence of such extensions show that, if there exists a realizable measure for a set of correlations there will in general be an infinite, in fact uncountable number of realizations.

Gibbs measures

The next question we ask is whether a specified set of correlation functions ρ_j, $j = 1, \ldots, k$, which can be realized by at least one point process, can also be realized by a Gibbs measure involving at most k-particle potentials. We show that for the case in which our system lives on a finite set Λ, e.g., a subset of the lattice this is possible whenever $\rho_1(\mathbf{x})$ satisfies $\rho_1(\mathbf{x}) < \bar{\rho}_1(\mathbf{x})$ for all $\mathbf{x} \in \Lambda$. We call such a measure k-Gibbsian.

The key ingredient in the argument is the fact that Gibbs measures are those which maximize the Gibbs-Shannon entropy of the measure μ,

$$S(\mu) \equiv -\sum_{\underline{\eta}} \mu(\underline{\eta}) \log \mu(\underline{\eta}) \tag{17}$$

subject to some specified constraints. In particular, we use the method of Lagrange multipliers to find a measure which maximizes the entropy, subject to the constraint of a given ρ_1 and ρ_2, then the maximizing measure will be 2-Gibbsian and the Lagrange multipliers obtained in this way will be the desired one body and pair potentials. The requirement that $\rho_1(\mathbf{x}) < \bar{\rho}_1(x)$ assures us that we are in the "interior" of the permissible set of measures and so the method of Lagrange multipliers is applicable.

The extension of the result to $\mathbb{Z}^d$ when ρ_1 and ρ_2 are translation invariant and to continuum systems is something we are still working on. The problem here is that we have an infinite number of constraints while the cases treated in the literature involve only a finite number of constraints.

A simple example

Determine for which densities ρ there exists a translation invariant point process on $\mathbb{R}^d$ with

$$g(\mathbf{r}) = \begin{cases} 0, & \text{if } |\mathbf{r}| \leq 1, \\ 1, & \text{if } |\mathbf{r}| > 1. \end{cases} \tag{18}$$

Condition (8) implies that (ρ, g) can only be realized if $\rho \leq (v_d 2^d)^{-1}$, where v_d is the volume of the ball with diameter 1 in $\mathbb{R}^d$ ($v_1 = 1$, $v_2 = \pi/4$, etc.).

In the other direction, (14) implies that for general d these correlations are indeed realizable if $\rho \leq e^{-1} v_d^{-1} 2^{-d}$. Thus the maximum density $\bar{\rho}(d)$ for which g is realizable satisfies

$$e^{-1} \leq 2^d v_d \bar{\rho}(d) \leq 1. \tag{19}$$

In one dimension: a simple construction shows realizability by a renewal process if $\rho \leq 1/e$. More complicated analysis gives

$$0.395 \leq \bar{\rho}(1) < 0.5$$

The gap between these upper and lower bounds remains as a challenge to further rigorous analysis.

On $\mathbb{Z}$ with nearest neighbor exclusion

$$\overline{\rho} \geq \frac{1}{4} = \max_{\rho \in [0,1]} \rho(1-\rho)$$

Proof: Start with Bernoulli measure with density ρ then eliminate any occupied site whose nearest neighbor to the right is also occupied.

III. Extension of Measures

Suppose that instead of being given translation invariant low order correlations on all of $\Omega = \mathbb{Z}^d$ ($\mathbb{R}^d$), we are given a measure μ on some finite subset $\Lambda \subset \Omega$. We then ask: can this measure be extended to a translation invariant μ on all of Ω?

Clearly a necessary requirement is that the marginal of μ_Λ on any subset $A \subset \Lambda$, be the same as that of $B = TA$, a translate of A, whenever $TA \subset \Lambda$. We call this property of μ_Λ, pre-translation invariance (PTI).

Remark

It is sometimes convenient to use spin notation rather than particle (i.e., lattice gas) notation in describing configurations in $\Lambda \subset \mathbb{Z}$. As usual if η is a particle variable taking values in $\{0, 1\}$ we introduce a corresponding spin variable $\sigma = 2\eta - 1$ taking values in $\{+1, -1\}$. This is convenient in particular because it permits us to write a measure μ_Λ directly in terms of the corresponding spin correslations:

$$\mu_k(\underline{\sigma}_\Lambda) = 2^{-|\Lambda|}(1 + \sum_{A \subset \Lambda, A \neq 0} \langle \sigma_A \rangle \sigma_A),$$

where as usual $\sigma_A = \prod_{i \in A} \sigma_i$ and the spin correlation $\langle \sigma_A \rangle$ denotes the expectation $\mu_\Lambda(\sigma_A)$ of σ_A in the measure μ_Λ.

One may think of the expectations $\langle\sigma_A\rangle$ as parameters which determine the measure; the condition that μ_Λ be PTI is simply that $\langle\sigma_B\rangle = \langle\sigma_A\rangle$ whenever B is a translate of A, with $A, B \subset \Lambda$. Note, however, that in using the above form to construct a measure μ_Λ with given $\langle\sigma_A\rangle$ one must check that it assigns a nonnegative probability to each configuration, $\underline{\sigma}$ in Λ.

Assume now that μ_Λ is indeed PTI. Then, the possibility of extension depends on the dimension. We show by explicit construction, for the case $\Omega = \mathbb{Z}$ (and expect also for $\mathbb{R}$) that all PTI measures are extendable. In higher dimensions, on the contrary, it is possible to construct PTI measures, e.g. on the unit square in $\mathbb{Z}^2$, which are not extendable.

Let me describe our explicit construction in $\mathbb{Z}$, where $\Lambda = \{0, 1, \ldots, k\}$ and $\mu_\Lambda = \mu_k(\eta_0, \ldots, \eta_k)$, $\eta_j = \{0, 1\}$ $\{\eta_0, \ldots, \eta_k\} \in X_k = \{0, 1\}^{k+1}$.

We define μ_{k-1} on X_{k-1} to be the marginal of μ_k:

$$\begin{aligned} \mu_{k-1}(\eta_0, \ldots, \eta_{k-1}) &= \sum_{\xi=0,1} \mu_k(\xi, \eta_0, \ldots, \eta_{k-1}) \\ &= \sum_{\xi=0,1} \mu_k(\eta_0, \ldots, \eta_{k-1}, \xi), \quad (20) \end{aligned}$$

where the equality of the two expressions in (20) expresses the PTI property of μ_k.

For $(\eta_0, \ldots, \eta_{k+1}) \in X_{k+1}$ one defines

$$\mu_{k+1}(\eta_0, \eta_1, \ldots, \eta_{k+1}) = \frac{\mu_k(\eta_0, \eta_1, \ldots, \eta_k)\mu_k(\eta_1, \eta_2, \ldots, \eta_{k+1})}{\mu_{k-1}(\eta_1, \ldots, \eta_k)} \tag{21}$$

Checking that μ is PTI and gives the correct marginal on X_k is straightforward.

It is now easy to see, by repeated application of the construction that if μ_k is a PTI measure on X_k then there exists a TI measure μ on X which extends μ_k.

Another way of writing (21) is
$\mu_{k+1}(\eta_{k+1}|\eta_0, \eta_1, \ldots \eta_k) = \mu_k(\eta_{k+1}|\eta_1, \ldots \eta_k)$

In fact for any $j > k$, we get

$$\mu_j(\eta_0, \eta_1, \ldots, \eta_j) = \mu_k(\eta_0, \ldots \eta_k) \prod_{i=1}^{j-k} \mu_k(\eta_{i+k} \mid \eta_i, \ldots \eta_{i+k-1}). \tag{22}$$

Eq. (22) says that we may regard the extension procedure as defining a Markov chain having finite memory, for which the transition probabilities depend on states at the previous k time steps. The TI extension μ of μ_k is then just the invariant measure on sample paths for this chain.

Maximal Entropy Extension: Gibbs Measures

The construction we have given for extending a PTI μ_k to μ_{k+1} is, as you may have recognized already, one that maximizes the Gibbs-Shannon entropy of the measure μ_{k+1}.

$$S(k+1) = -\sum \mu_{k+1}(\eta_0, \ldots, \eta_{k+1}) \log \mu_{k+1}(\eta_0, \ldots, \eta_{k+1}) \tag{23}$$

subject to the constraints that the projection of μ_{k+1} on the $\{\eta_0, \ldots \eta_k\}$ and on the set $\{\eta_1, \ldots \eta_{k+1}\}$ be given.

This follows from the general fact that given a measure $\mu(A, B)$ on $A \cup B$ and another measure $\mu(B, C)$ on $B \cup C$ which agree in their projections on B, then any extension to a measure on $A \cup B \cup C$, $\mu(A, B, C)$ which agrees with $\mu(A, B)$ and $\mu(A, C)$ has the property that

$$S(A, B, C) \leq S(A, B) + S(B, C) - S(B) \tag{24}$$

Equality in (24) is achieved when

$$\mu(A, B, C) = \mu(A, B)\mu(B, C)/\mu(B) \tag{25}$$

(with the obvious caveat that if the denominator vanishes so does the numerator) This is exactly the form of extension we have made from μ_k to μ_{k+1}. With some abuse of notation:

$$A = \{0\}, \quad B = \{1, \ldots, k\}, \quad C = \{k + 1\} \tag{26}$$

It is also easy to see that the translation invariant measure we get this way will be “Gibbs” with “interaction potentials” which involve at most $k+1$ sites, i.e., they will have a range at most k. The interesting part is that this entropy maximizing extension is automatically TI on $\mathbb{Z}$.

Since entropy maximizing measures are “generally” Gibbsian we have that, in $d = 1$ the restriction of a Gibbs measure with TI length k-interactions on $j > k$ sites is the entropy maximizing PTI measure μ_j obtained from μ_k by our construction.

Our one dimensional construction also gives an expression for the entropy of any block of length $k+j$ obtained from a translation invariant Gibbs measure on $\mathbb{Z}$ with interaction of range k or less. This is just

$$S(k+j) = S(k) + j[S(k) - S(k-1)] \qquad (27)$$

This implies that the specific entropy, $s(\mu)$, i.e., the entropy per site of the TI maximal entropy extension is equal to $[S(k) - S(k-1)]$.

When $S(k) = S(k-1)$ then one can show that the measure μ lives on periodic configuration of length L, $L \leq 2^k$.

To see this note that $S(k) = S(k-1)$ implies that the configuration at site k is fully determined by the configuration at sites $\{0, 1 \ldots, k-1\}$ which means that the process is deterministic after k sites. This implies that there are at most only 2^k configurations with non zero probability.

Remark

Everything said above extends immediately to measures defined on S^{Λ_k}. As an example, take $S = \{0,1\}^{\Lambda_m}$ so that S^{Λ_k} may be identified with the space $\{0,1\}^{\Lambda_k \times \Lambda_m}$ of particle configurations on a $k \times m$ rectangle. If such a measure is PTI under translations in the first component then it may be extended to a measure on configurations on $\Lambda \times \Lambda_m$, where Λ is any interval containing Λ_k, and hence to a measure on configurations on $\mathbb{Z} \times \Lambda_m$. Unfortunately, however, this extension procedure need not maintain the PTI property for translations in the second component.

Thus if we start with a measure on a strip of $\mathbb{Z}^2$ of height h and length k which is PTI in both directions we can extend it to either an infinite horizontal or vertical measures, but not necessarily to a TI measure in both directions.

An example of a non-extendable PTI in $\mathbb{Z}^2$ is to take for Λ the unit square and assign the following probabilities to the 16 possible configurations: probability 1/4 to the following four configurations

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

and probability zero to all other (12) configurations. It is simple to check that this increase is PTI but has <u>no</u> PTI extension to a 3 by 3 square.

What is still true however is the following (still needs some checking for precise condition). Given a TI measure μ on $\mathbb{Z}^d$ there exists a unique Gibbs measure μ_G on $\mathbb{Z}^d$ with TI interactions of range less than or equal to L such that the projection of μ and μ_G agree on any set $\Lambda \subset \mathbb{Z}^d$, with diameter of Λ, $D(\Lambda) \leq L$.

Periodic Extensions

It is possible to show that among the TI extensions of PTI μ_k there will always be measures supported on periodic configurations of finite length. These measures are superpositions of measures concentrated on "minimal" periodic configurations of length p. They give weight $1/p$ to each of the p translates of the p-periodic configurations.

Example: Suppose $\mu_1(\eta_0, \eta_1)$ gives weight $\alpha_{1,1}$ to $(1,1)$, $\alpha_{0,0}$ to $(0,0)$ and $\alpha_{0,1} = (1 - \alpha_{1,1} - \alpha_{0,0})/2$ to $(1,0)$ and to $(0,1)$, then we can obtain a TI measure by giving weight $\alpha_{1,1}$ to configurations $(\ldots, 1, 1, 1, \ldots)$, $\alpha_{0,0}$ to $(\ldots, 0, 0, 0, \ldots)$, $\alpha_{0,1}$ to those which consist of $(\ldots, 0, 1, 0, 1, \ldots)$ and $\alpha_{1,0} = \alpha_{0,1}$ to $(\ldots, 1, 0, 1, 0, \ldots)$.

The proof of this statement for general k is based on the use of De Bruijn graphs.

The (binary) *De Bruijn graph of order* k (or sometimes *dimension* k), G_k, is the directed graph with 2^k vertices and 2^{k+1} edges, labeled respectively by all binary strings of length k and length $k+1$, in which for any binary digits a and b and binary string θ of length $k-1$ the edge $a\theta b$ runs from vertex $a\theta$ to vertex θb (see Figure 1).

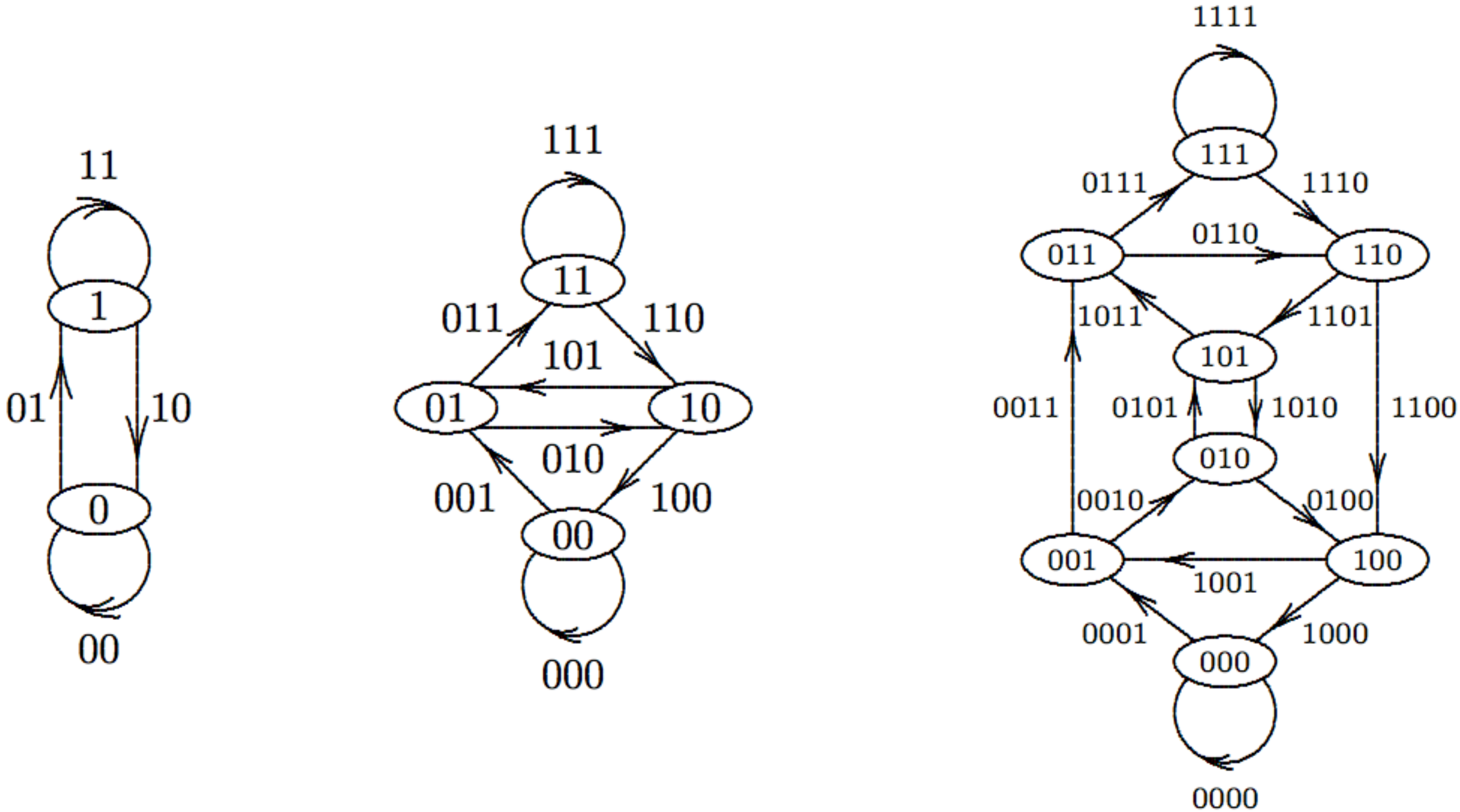

Figure 1: The first three De Bruijn graphs

Note in particular that G_k contains two loops, labeled respectively by $00\ldots0$ and $11\ldots1$, but no multiple edges. Since the edges of G_k are labled by the elements of X_k, it is clear that any probability measure μ_k on X_k corresponds to an assignment of a nonnegative *current* j_η to each edge η of G_k in such a way that $\sum_\eta j_\eta = 1$; the correspondence is of course via $j_\eta = \mu_k(\{\eta\})$. The terminology "current" is appropriate because one checks easily that μ_k is PTI if and only if the current is conserved at each vertex of G_k, that is, if for each vertex ξ the sum of the currents on the two edges entering ξ is equal to the sum of the currents on the two edges leaving ξ.

Suppose that $\mathcal{P}$ is a closed path in G_k, that is, a sequence of $|\mathcal{P}|$ edges $\eta^{(1)}, \ldots, \eta^{(|\mathcal{P}|)}$ in that order (possibly with repetitions), and that $\mathcal{P}$ is *minimal* in the sense that there is no shorter path $\mathcal{P}'$ such that $\mathcal{P}$ is obtained by tracing $\mathcal{P}'$ several times. With $\eta^j = \eta^j_0 \ldots \eta^j_k$ we let $\nu_{\mathcal{P}}$ denote the measure corresponding to the periodic configuration $\ldots \eta_0^{(|\mathcal{P}|)} \eta_0^{(1)} \ldots \eta_0^{(|\mathcal{P}|)} \eta_0^{(1)} \cdots \in X$; this is clearly a bijective correspondence between minimal closed paths and measures on "primitive" periodic configurations.

It follows then that every PTI measure μ_k on X_k is the marginal of a TI measure ν with ν a (finite) convex combination of the k-primitive periodic measures $\nu_{\mathcal{C}}$. In particular, every such μ_k has an extension to a TI measure supported on periodic configurations of finite length.

Entropy Minimizing Measures

It seems natural to ask which of the TI extensions of μ_k minimizes the Gibbs-Shannon entropy. It is clear from the above that these measures, which due to the concavity of the entropy will be at the boundary of the permitted set and need not be unique, will have finite entropy. Indeed we can prove that any entropy minimizing measure will be a finite superposition of the periodic measures discussed above.

IV. Density Matrices

The question we ask for quantum systems is similar to the one we asked for classical systems: Given density matrices $\rho(1,2)$ and $\rho(2,3)$ on the Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\mathcal{H}_2 \otimes \mathcal{H}_3$ such that

$$\mathrm{tr}_1 \rho(1,2) = \mathrm{tr}_3 \rho(2,3) = \rho(2),$$

is there a density matrix $\rho(1,2,3)$ acting on $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ such that

$$\mathrm{tr}_1 \rho(1,2,3) = \rho(2,3), \quad \mathrm{tr}_3 \rho(1,2,3) = \rho(1,2)$$

Remember that in the classical case this was always possible: we simply set $\mu(1,2,3) = \mu(1,2)\mu(2,3)/\mu(2)$ and that this gave in fact the maximal entropy extension.

This no longer works for density matrices due to entanglement. It is easy to give examples, e.g., suppose both $\rho(1,2)$ and $\rho(2,3)$ correspond to pure states (which are not themselves direct products of pure states for $\rho(1), \rho(2), \rho(3)$) then there is no way to extend $\rho(1,2)$ and $\rho(2,3)$ to $\rho(1,2,3)$. An easy way to see this is to note that the strong subadditivity for the von Neumann entropy, $S(\rho) = -\mathrm{tr}\rho \log \rho$
says that just like in the classical case

$$S(\rho_{123}) \leq S(\rho_{12}) + S(\rho_{23}) - S(\rho_2) \tag{28}$$

so if $S(\rho_{12}) = S(\rho_{23}) = 0$, the inequality cannot hold unless $S(\rho_2) = 0$.

The inequality (28)goes under the name of strong subadditivity (SSA) and was proven by Lieb and Ruskai in the 70's. It is a hard theorem and is of central importance in quantum information.
The only positive result we have in the direction of extension is when both $\rho(1,2)$ and $\rho(2,3)$ are a convex combination of direct products

$$\left.\begin{aligned} \rho(1,2) &= \sum_j \lambda_j \sigma_1^{(j)} \otimes \sigma_2^{(j)} \\ \rho(2,3) &= \sum \lambda_j \sigma_2^{(j)} \otimes \sigma_3^{(j)} \end{aligned}\right\} \quad \lambda_j > 0,\ \sum_j \lambda_j = 1$$

with the $\sigma_2^{(j)}$ the same in both representations.

When there exists an extension to $\rho(1,2,3)$ there will also be a maximal entropy extension by concavity of the entropy. It seems however that even in that case the maximal entropy extension will in general not saturate the SSA to an equality.

This means that all the relations we derived for the entropies of Gibbs measures on $\mathbb{Z}$ in the classical case do not extend in general to quantum systems.

So to conclude: what, if any, are the necessary and sufficient conditions for extension of density matrices. A solution gets a bottle of good wine and a list of open problems.

THANK YOU