

Positive representations of complex distributions on groups

L. L. Salcedo*

*Departamento de Física Atómica, Molecular y Nuclear and
Instituto Carlos I de Física Teórica y Computacional,
Universidad de Granada, E-18071 Granada, Spain.*

(Dated: May 7, 2018)

A normalizable complex distribution $P(x)$ on a manifold \mathcal{M} can be regarded as a complex weight, thereby allowing to define expectation values of observables $A(x)$ defined on \mathcal{M} . Straightforward importance sampling, $x \sim P$, is not available for non positive P , leading to the well-known sign (or phase) problem. A positive representation $\rho(z)$ of $P(x)$ is any normalizable positive distribution on the complexified manifold \mathcal{M}^c , such that, $\langle A(x) \rangle_P = \langle A(z) \rangle_\rho$ for a dense set of observables, where $A(z)$ stands for the analytically continued function on \mathcal{M}^c . Such representations allow to carry out Monte Carlo calculations to obtain estimates of $\langle A(x) \rangle_P$, through the sampling $z \sim \rho$. In the present work we tackle the problem of constructing positive representations for complex weights defined on manifolds of compact Lie groups, both Abelian and non Abelian, as required in lattice gauge field theories. Since the variance of the estimates increase for broad representations, special attention is put on the question of localization of the support of the representations.

Contents

I. INTRODUCTION

I. Introduction	1
II. Representations of complex probabilities	2
A. Definition of representation	2
B. Existence of positive representations	2
C. Localization of the support of positive representations	3
III. Localized representations of Abelian groups	4
A. Two-branch representations in one-dimension	4
B. Two-branch representations in higher dimensions	5
1. Strict two-branch approach	5
2. Uniform two-branch approach	6
C. Representations from convolutions	7
D. Complex representations and linearity	7
IV. Localized representations on Lie groups	8
A. Representations on groups	9
B. Two-branch representations	9
C. An $SU(2)$ example	11
D. Representations through convolutions	12
E. Representations in matrix groups	12
V. Removal of singular kernels and examples	13
A. Singular kernels	13
B. Case study I	15
C. Case study II	16

VI. Summary and conclusions

Acknowledgments

References

Many a scientific problem, in physics or otherwise, can be reduced to obtaining the expectation values of observables, assigning a weight to each existing configuration of some system. When the number of configurations is large, a Monte Carlo sampling method is often the best option, or even the only available one in practice [1]. However, the route through importance sampling is blocked when the weights are not definite positive. This constitutes the well-known sign problem [2].

The sign (or phase) problem arises in many context including statistical mechanics, condensed matter, nuclear physics and quantum field theory, often related to the presence of fermions in many body systems. In the context of lattice gauge field theory the problem arises, for instance in attempting to study QCD at finite baryonic density. The impediment is that in the Euclidean formulation the Boltzmann weight is reflection positive, as required by unitarity [3], but not directly positive in the presence of a chemical potential [4].

Several techniques have been tried to solve or soften the sign problem [5]. Among the potentially exact ones, one approach is that of reweighting, that is, applying Monte Carlo by sampling a suitable positive distribution and including the ratio of weights as a factor in the observable. The method is correct and rigorous but it suffers from the well-known overlap problem: even for seemingly similar weights, differences increase exponentially with the size of the system. As a consequence variances in the estimates increase and the signal-to-noise ratio becomes negligible [6].

Another technique aiming at solving the problem exploits the analyticity of the complex weight in many practical cases, including lattice gauge field theory. Actually analyticity is routinely used to go from Lorentzian to Euclidean metrics in those settings. The complex Langevin equation approach [7, 8] simply applies the stochastic Langevin equation to the complex case relying on the good analytical properties of the action, and observables are computed through their analytical extension. This elegant approach enjoys nice features, above all, that of preserving the locality of the standard Monte Carlo

*Electronic address: salcedo@ugr.es

algorithms, and has been successfully applied to some practical problems [9–11]. Regrettably, the technique is not mathematically robust. Even in simple one-degree-of-freedom systems the algorithm may not converge, or converge to unwanted solutions [12–15]. A recent review of the present status of the complex Langevin technique can be found in [16].

A more recently introduced approach to cope with the sign problem is that of Lefschetz thimbles [17, 18]. It also relies on analytical continuation of the action and the observables, using an optimal deformation of the original real manifold and an additional residual reweighting. The need of several submanifolds (thimbles), with unknown relative complex weights, hinders a straightforward application of the method, which is very promising [19].

The complex Langevin approach aims at constructing a real and positive distribution on the complexified manifold, in such a way that the expectation values of the analytically continued observables correctly reproduce the expectation values of the original complex weight defined on the real manifold of configurations of the system. Such a real and positive distribution, whether originated from complex Langevin or not, was called a *representation* (of the complex weight) in [20].

The explicit construction of direct representations (i.e., constructed without a complex Langevin approach) was undertaken in [20]. The existence of positive representations for one-dimensional complex weights was established in [21], and for very general complex weights and manifolds in [22]. Further constructions have been presented in [23–30].

The two-branch approach in [22, 25, 27, 28, 31] is particularly suitable in order to obtain localized representations. This is a major issue in the representation approach since there is an overlap problem, similar to that of reweighting, related to the extension of the representation, which reflects on the variance of the Monte Carlo estimates. Such an approach has been applied in [25] to carry out a Monte Carlo sampling with a complex version of the heat bath method.

Previous works have dealt mainly with complex weights defined on manifolds of Abelian groups, \mathbb{R}^n or $U(1)^{\times n}$. The case of non Abelian groups is needed in practical applications, such a lattice gauge field theory. This case was treated in [22] in a rather formal way, showing existence constructively. In the present work we address the issue of finding explicit direct representations of complex weights defined on non Abelian matrix groups. The main concepts are revised in Sec. II. After a review of the two-branch approach in $U(1)^{\times n}$, we present an improved prescription to symmetrically treat all the variables, in the many-dimensional case in Sec. III. The case of compact non Abelian Lie groups is considered in Sec. IV, where formulas are derived for matrix groups, formally applying to the non compact case too. Obstructions arise in our approach when some group representations contain singlet subrepresentations, with respect to the subgroup generated by the element making the lifting to the complex manifold. This issue is dealt with in Sec. V, and also some examples are analyzed in detail. Sec. VI summarizes our conclusions.

II. REPRESENTATIONS OF COMPLEX PROBABILITIES

A. Definition of representation

We consider continuous degrees of freedom throughout. Let $P(x)$ be a complex distribution defined on some manifold \mathcal{M} . In applications, $P(x) = e^{-S(x)}$ where $S(x)$ is the action of the system with configuration x . We assume that P has a non vanishing normalization, $\int d\mu(x) P(x) \neq 0$. With some abuse of language, we will refer to P as a *complex probability*, because expectation values of observables $A(x)$ can be defined with the same rules as for ordinary (real and positive) probability densities, i.e.,

$$\langle A \rangle_P = \frac{\int d\mu(x) P(x) A(x)}{\int d\mu(x) P(x)}, \quad (2.1)$$

where $d\mu(x)$ is a suitable positive measure on \mathcal{M} .

Unfortunately, when $P(x)$ is not positive definite, importance sampling, $x \sim P(x)$, is meaningless and this prevents the straightforward application of a Monte Carlo method. This is the well-known sign problem.

Ever since the conception of the complex Langevin algorithm [7, 8], one of the approaches devised to sort out this impediment is to replace the original manifold by its complexified version \mathcal{M}^c , the observables by their holomorphic extension, $A(z)$, and the complex probability by an ordinary probability distribution $\rho(z)$ defined on \mathcal{M}^c .¹ An obvious condition on ρ is

$$\int d\mu(x) P(x) A(x) = \int d\mu^c(z) \rho(z) A(z) \quad \text{for all } A. \quad (2.2)$$

In this case, we say that ρ is a *representation* of the complex probability P . This property implies

$$\langle A(x) \rangle_P = \langle A(z) \rangle_\rho, \quad (2.3)$$

hence estimates obtained from ρ reproduce those of P . An additional condition is $\rho(z) \geq 0$, so that importance sampling can be applied to ρ , therefore we aim at *positive representations* of complex probabilities.

Regarding Eq. (2.2), let us remark that the condition can be relaxed by allowing a different normalization in P and ρ , and also *for all* A really means a suitable (ideally dense with respect some topology) set of test functions, as in standard distribution theory.

B. Existence of positive representations

Obviously, representations exists for any $P(x)$, for instance $\rho(z) = P(x) \delta(y)$, where y denotes the coordinates in the imaginary direction in \mathcal{M}^c . Less trivially, *positive* representations

¹ In this work, following [14, 15, 20, 22, 25, 31], P denotes the complex density defined on the real manifold, while ρ denotes the real density defined on the complex manifold. The notation exchanging the roles of the symbols P and ρ is also frequently used in the literature [13, 16, 27, 32, 33].

also exist for very general complex probabilities [22], and the solution is by no means unique. The non uniqueness follows from the fact that the set of holomorphic observables constraining ρ is only a subset of all test functions on the complexified manifold.

An explicit construction for $\mathcal{M} = \mathbb{R}^n$ has been given in [22], as follows: Without loss of generality let P be normalized and let P_0 be a positive probability and also normalized,

$$1 = \int d^n x P(x) = \int d^n x P_0(x), \quad P_0(x) \geq 0. \quad (2.4)$$

Then $P - P_0$ integrates to zero and can be written as a derivative:

$$P(x) = P_0(x) + \nabla \cdot (P_0(x) \mathbf{H}(x)), \quad (2.5)$$

where the vector field $\mathbf{H}(x)$ can be chosen in many ways. Equivalently,

$$P(x) = \int d^n x' P_0(x') (\delta(x - x') + \mathbf{H}(x') \cdot (\nabla \delta)(x - x')). \quad (2.6)$$

So it is sufficient to obtain positive representations of distributions of the type $\delta(x) + \mathbf{h} \cdot \nabla \delta(x)$ and weight them with P_0 . An explicit solution is

$$\rho(z) = \int d^n x' P_0(x') \int d^2 z' q(z') \delta(z - x' - z' \mathbf{H}(x')), \quad (2.7)$$

where $q(z)$ is any positive representation of $\delta(x) + \delta'(x)$ (a one-dimensional distribution), such as

$$q(z) = \frac{1}{8\pi} \left| 1 - \frac{z}{2} \right|^2 e^{-|z|^2/4}. \quad (2.8)$$

That is, the average of $A(x - z \mathbf{H}(x))$, with $x \sim P_0$ and $z \sim q$, correctly reproduces $\langle A \rangle_P$. The analogous construction for arbitrary compact Lie groups has been given in [22].

C. Localization of the support of positive representations

While the problem of finding positive representations of generic complex distributions is formally solved, the impediments for systems of large dimensionality remain in practice. Indeed, the vector field \mathbf{H} is not easy to obtain. Even more importantly, in general, the magnitude of \mathbf{H} will scale as $|P|$ as the number of degrees of freedom (or volume) increases. Since the action S scales as the volume, this implies an exponential growth in \mathbf{H} which in turn entails an exponential growth in the size of the support of the representation ρ and so in the dispersion of the random variable z in $A(z)$. This would translate into an exponentially large variance in the Monte Carlo estimates.

This is an important point of the representation approach: in the standard case of positive probabilities, the sampling $x \sim P$ is uniquely defined by P .² This is no longer true when the

estimate is obtained by means of a representation since many different representations exist. These are all formally equivalent (as all of them fulfill Eq. (2.2)) but they can be very different regarding the variance of the estimates obtained from them. Ideally one would like a ρ with a support as localized as possible in order to reduce the dispersion. This problem is analogous to that in the reweighting approach, where a maximum overlap is desirable. A complete overlap is not possible if P is complex, and also in the representation approach a perfect localization of ρ on the real manifold is not attainable.

Since observables tend to grow wildly as one departs from the real manifold, representations close to it are preferable in general. The *width* of a representation ρ can be defined as the size of its support in the imaginary direction, and for a given complex probability there are bounds on how narrow any positive representation of it can be. As one would expect, the more complex (in the sense of less positive definite) a complex probability is the wider is its narrowest positive representation. Not surprisingly, obtaining wider (and so worse quality) representations poses no problem.³

Regarding localization of the support of any positive representation of a given complex probability, a general observation is made [25]: for any observable A , the support of ρ must contain values of $|A(z)|$ larger than $|\langle A \rangle_P|$ (note that this quantity is independent of the choice of ρ).⁴ In particular a concrete bound follows (in the one-dimensional case but can be extended to higher dimensions), namely, if the support of $\rho(z)$ is entirely contained in a horizontal strip $Y_2 < y < Y_1$, one must have

$$Y_1 \geq \max_{k>0} \left(\frac{1}{k} \log |\tilde{P}(k)| \right), \quad Y_2 \leq \min_{k<0} \left(\frac{1}{k} \log |\tilde{P}(k)| \right), \quad (2.9)$$

where $\tilde{P}(k)$ is the Fourier transform of $P(x)$. In practice, these bounds are quite tight for typical P 's [25].

Many additional conditions can be imposed on the support of a positive measure ρ representing a complex probability P . For instance, for any observable $A(x)$, let $a \equiv \langle A \rangle_P$, and let the two nonempty complementary regions in \mathcal{M}^c be defined by $\mathcal{A}_> = \{\text{Re}(A(z)) \geq \text{Re}(a)\}$ and $\mathcal{A}_< = \{\text{Re}(A(z)) < \text{Re}(a)\}$ (we exclude the trivial case of a constant A). Then the relation

$$\langle \text{Re}(A) \rangle_\rho = \text{Re}(a) \quad (2.10)$$

requires that the support of ρ must have some overlap with both regions as it cannot be *entirely* contained in any of them. The fulfillment of this condition for all observables A puts

reduce the variance, is of academic interest only, first because sampling is expensive and many observables are to be considered, and second because $P = e^{-S}$ behaves exponentially with respect to typical observables (including S) and so sampling P is mandatory.

³ Applying an isotropic diffusion process to any positive representation produces another, less localized, positive representation of the same complex probability [20].

⁴ This simple consideration, for instance, rules out that the complex Langevin algorithm could produce a proper representation for the action $S(x) = x^4/8 + 2ix$ [25].

² The influence of the concrete observable A on the sampling, in order to

constraints on the allowed support of positive (or more generally real) representations. Of course, taking $e^{-i\theta}A$, the same consideration holds for $\text{Re}_\theta(A) \equiv \text{Re}(e^{-i\theta}A)$ ($\theta \in \mathbb{R}$), and for $\text{Im}(A)$ in particular.

The usefulness of this kind of relations can be seen in the following example. Let $P(x) = e^{-s(x)}$ with $S(x) = x^4 - 2x^2 - 2ix$. For this complex probability $\langle x \rangle = -7.83i$. Since this value is below the real axis, any positive $\rho(z)$ representing $P(x)$ must have some support below the real axis. However, if one applies a standard complex Langevin prescription, the stationary solution for ρ will be above the real axis: the velocity drift points upwards along the real axis so the complex Langevin walker can never cross the real axis once she is above it. We have proven failure of complex Langevin in this case without an explicit simulation of the stochastic process.

Summarizing, positive representation exists for arbitrary or very general complex probabilities, and localized representations are highly preferable from the point of view of Monte Carlo calculations. It is also noteworthy that one can impose on the representations the same symmetries enjoyed by the complex probability itself (provided the symmetrization procedure is compatible with the analytical extension, which is often, if not always, the case).

III. LOCALIZED REPRESENTATIONS OF ABELIAN GROUPS

The complex probabilities considered in this section are defined on \mathbb{R}^n or periodic versions of it, so they can be viewed as complex probabilities on Abelian groups, namely, $(\mathbb{R}^n, +)$ or $U(1)^{\times n}$ or mixed cases of them.

We first review the construction of localized representations carried out in [25]. A similar construction has been derived independently by Seiler and Wosiek in [27]. The one-dimensional and higher dimensional cases are discussed. Subsequently, a more systematic and satisfactory treatment of the higher dimensional case is introduced.

An important feature of the representations discussed here is that their support is composed of (a finite number of) parallel copies of the real manifold, at different heights in the imaginary direction. Therefore, these representations can be used with any holomorphic observable, regardless of how wildly such observable may behave in the deep imaginary region. Analogous constructions will be obtained for complex measures defined on more general groups in the next section.

A. Two-branch representations in one-dimension

Consider $P(x)$ defined on $U(1)$ (the case $x \in \mathbb{R}$ is completely analogous in most respects and is described in [25]). We use the normalization

$$1 = \int_0^{2\pi} \frac{dx}{2\pi} P(x) \quad (3.1)$$

(and assume P to be normalized throughout the construction).

A suitable set of test functions is e^{-ikx} , hence we aim at finding a positive representation $\rho(z)$ such that ($z = x + iy$)

$$\int_0^{2\pi} \frac{dx}{2\pi} P(x) e^{-ikx} = \int_0^{+\infty} \frac{dy}{2\pi} \rho(z) e^{-ikz}, \quad \forall k \in \mathbb{Z}. \quad (3.2)$$

As said, there are many solutions for ρ . The (symmetric) *two-branch* solution is of the form

$$\rho(z) = Q_+(x)\delta(y-Y) + Q_-(x)\delta(y+Y) \quad (3.3)$$

($Y \neq 0$). That is, ρ has support on the two horizontal lines $y = \pm Y$, parallel to the real axis. Each of the two branches is a copy of the real manifold. The width is $2Y$ and this is a parameter to be chosen in the construction (without loss of generality we can assume $Y > 0$).

Such $\rho(z)$ will be a representation of $P(x)$ provided

$$P(x) = Q_+(x - iY) + Q_-(x + iY) \quad (3.4)$$

(in the sense of analytical extension). Indeed, it is readily verified that this equation implies

$$\langle A(x) \rangle_P = \sum_{\sigma=\pm} N_\sigma \langle A(x + i\sigma Y) \rangle_{Q_\sigma}, \quad N_\pm \equiv \int \frac{dx}{2\pi} Q_\pm(x). \quad (3.5)$$

The two functions $Q_\pm(x)$ are (almost) uniquely determined by the requirement of being *real* (for real x). Using this condition in Eq. (3.4) yields, for the Fourier modes,

$$\tilde{Q}_{\pm,k} = \pm \frac{e^{\pm kY} \tilde{P}_k - e^{\mp kY} \tilde{P}_{-k}^*}{2 \sinh(2kY)} \quad (k \neq 0), \quad (3.6)$$

where

$$P(x) = \sum_k \tilde{P}_k e^{ikx}, \quad Q_\pm(x) = \sum_k \tilde{Q}_{\pm,k} e^{ikx}. \quad (3.7)$$

The *zero modes* of these functions, $\tilde{Q}_{\pm,0} = N_\pm$, are not fixed by the condition of ρ being a representation of P . Obviously, a constant term can be freely added to Q_+ and subtracted to Q_- without violating Eq. (3.4). The normalization of P only implies the combined relation

$$N_+ + N_- = 1. \quad (3.8)$$

By construction $Q_\pm(x)$ are real for any value of Y . In general they are not positive definite and diverge for small Y , except when $P(x)$ is real. In that case

$$\tilde{Q}_{\pm,k} = \frac{\tilde{P}_k}{2 \cosh(kY)} \quad (\text{real } P(x)) \quad (3.9)$$

and $Q_\pm(x) \rightarrow \frac{1}{2}P(x)$ as $Y \rightarrow 0$.

Going in the opposite direction, inspection of eq. (3.6) shows that for *sufficiently large* Y , all non zero Fourier modes in Eq. (3.6) become arbitrarily small if \tilde{P}_k is exponentially bounded.⁵ Hence, taking $N_\pm > 0$, it follows that eventually

⁵ That is, if $|\tilde{P}_k| < K e^{Y|k|}$ for some $K, Y > 0$.

N_{\pm} dominate the Fourier sum and $Q_{\pm}(x)$ are guaranteed to be positive. This shows that essentially any periodic complex probability admits a positive representation of the two-branch type. Explicit examples of representations of the two-branch type can be found in [25].

As already noted, in practice it is desirable to have a width as small as possible. The prescription to achieve this has been given in [25]: starting from the bounds in Eq. (2.9), Y can be continuously increased. Eventually, for some critical value $Y = Y_c$

$$\min(q_+, q_-, q_+ + q_-) = -1, \quad q_{\pm} \equiv \min_x \sum_{k \neq 0} \tilde{Q}_{\pm,k} e^{ikx} = \min_x (Q_{\pm}(x) - N_{\pm}). \quad (3.10)$$

For $Y \geq Y_c$, suitable $0 \leq N_{\pm} \leq 1$ exist so that $Q_{\pm}(x)$ are positive for all x . In particular for $Y = Y_c$, $\min_x Q_{\pm}(x) = 0$.

The construction in \mathbb{R} (as opposed to $[0, 2\pi]$) is quite similar, the main difference being that the freedom in sharing zero modes between the two sheets $y = \pm Y$ no longer exists [25]. We discuss further the noncompact case at the end of Sec. III D.

It can be noted that we have chosen as support of our representation exactly two horizontal lines and equidistant from the real axis, $y = \pm Y$. As discussed in [25] an asymmetric choice is possible but in practice no substantial gain is achieved by doing that (for generic complex probabilities). So we favor simplicity in our construction in order to facilitate its extension to more complicated scenarios.

Another question is the use of more branches, $y = Y_1, \dots, Y_n$. Also nothing is gained in practice. Moreover, since one must impose positivity on each branch separately, this implies a larger number of conditions which translate into larger values of Y_i (and so larger variances). In [22] each Fourier mode $a_k e^{ikx}$ was treated separately. This is legitimate but not optimal. Since a single Fourier mode has zero normalization (except $k = 0$) one must share the total normalization of P (namely, 1) among the Fourier modes, and obtain a positive representation of each $n_k + a_k e^{ikx}$. For a fixed amplitude a_k , the smaller the normalization n_k , the wider the representation (larger Y). So the sharing among modes, $1 = \sum_k n_k$, must be optimized and even so, imposing positivity for the representation of each separate mode requires larger values of Y . The great advantage of the two-branch approach of [25] is that all the modes are added on the same branch (same support) and they compensate each other to have a positive function with a minimal common width.

B. Two-branch representations in higher dimensions

The above construction can be generalized to functions defined on the torus $[0, 2\pi]^n$, or equivalently $U(1)^{\times n}$, although this is not completely straightforward.

1. Strict two-branch approach

For normalized P , one can tentatively propose

$$\rho(z) = Q_+(x)\delta(y - Y) + Q_-(x)\delta(y + Y), \quad (3.11)$$

where the two functions $Q_{\pm}(x)$ are positive and the construction depends on the parameters $\mathbf{Y} = (Y_1, \dots, Y_n)$. The representation condition is equivalent to requiring

$$P(x) = Q_+(x - i\mathbf{Y}) + Q_-(x + i\mathbf{Y}), \quad (3.12)$$

and in terms of the Fourier modes this implies (demanding that $Q_{\pm}(x)$ should be real)

$$\tilde{Q}_{\pm,k} = \pm \frac{e^{\pm k \cdot \mathbf{Y}} \tilde{P}_k - e^{\mp k \cdot \mathbf{Y}} \tilde{P}_{-k}^*}{2 \sinh(2k \cdot \mathbf{Y})} \quad (k \cdot \mathbf{Y} \neq 0). \quad (3.13)$$

Note that $\tilde{Q}_{-,k}$ is just $\tilde{Q}_{+,k}$ with $-\mathbf{Y}$ instead of \mathbf{Y} .

Once again, the constant modes,

$$N_{\pm} = \int \frac{d^n x}{(2\pi)^n} Q_{\pm}(x), \quad (3.14)$$

are not fixed since, being constant under analytical extension, they can be moved freely between the two branches, Eq. (3.12). Also, for large enough \mathbf{Y} (assuming $k \cdot \mathbf{Y} \neq 0$) all non constant Fourier modes become small and the distributions $Q_{\pm}(x)$ eventually become positive for positive N_{\pm} .

Clearly the singular modes, i.e., those with $k \cdot \mathbf{Y} = 0$, pose a problem. This is for the same reason $k = 0$ is special: since Q_{\pm} are real, if one integrates over x on both sides of Eq. (3.12) the resulting equation is only consistent if the normalization of P is also real. Equivalently, the zero (constant) mode is unchanged by the shifts $x \rightarrow x \pm i\mathbf{Y}$ from the real to the complex manifold. By the same token, the singular modes with $k \cdot \mathbf{Y} = 0$ are not affected by the complex shift and the equation is only consistent if $P(x)$ happens to be real for those particular modes. For the zero mode, the reality condition is fulfilled due to our previous requirement that P should be normalized, but the remaining singular modes cannot be fixed by the same requirement.

An easy solution would be to take for the components of \mathbf{Y} suitable irrational numbers in such a way that the combination $\sum_i k_i Y_i$ can never be exactly zero (e.g. $\mathbf{Y} = (1, \sqrt{2}, \sqrt{3})$). However, such prescription is rather arbitrary and has several drawbacks: i) although $k \cdot \mathbf{Y}$ would not be exactly zero it could be arbitrarily small when many modes are relevant and this is numerically problematic.⁶ ii) As the problem worsens when all components of \mathbf{Y} are similar, this suggests using very dissimilar components. Unfortunately, positiveness of $Q_{\pm}(x)$ requires a sufficiently large vector \mathbf{Y} but too large values entail large variances; dissimilar values of the components of \mathbf{Y} imply that some of these components would be larger

⁶ In fact, in the noncompact case ($\mathcal{M} = \mathbb{R}^n$ rather than a torus) k is continuous and $k \cdot \mathbf{Y} = 0$ would not be avoided.

than necessary (to allow the shorter components to be sufficiently large). iii) Most importantly, if the various degrees of freedom represented by the variables $\mathbf{x} = (x^1, x^2, \dots, x^n)$ play a similar role in the action (a similarity that is often enforced by concrete symmetries of the action) one would request that \mathbf{Y} should also contain similar components for all of them, without ad hoc variation from one component to another, with no basis on the action or the physical problem at hand.

2. Uniform two-branch approach

A better solution is to use different displacement vectors \mathbf{Y} for different Fourier modes.⁷ Implicitly this implies to introduce further branches, i.e., further copies of the real manifold. In order to encompass the uniformity criterion noted above, in which all variables should play a similar role, a natural prescription is to introduce 2^n branches, a duplication for each degree of freedom. Each branch is characterized by a vector of n bits, $\boldsymbol{\sigma} = (\pm, \dots, \pm)$, so that

$$\mathbf{Y} = (\pm Y, \dots, \pm Y) = Y\boldsymbol{\sigma}. \quad (3.15)$$

Correspondingly, there are 2^n real and positive functions $Q_{\boldsymbol{\sigma}}(\mathbf{x})$ defined on the real manifold, and the representation condition becomes

$$P(\mathbf{x}) = \sum_{\boldsymbol{\sigma}=(\pm, \dots, \pm)} Q_{\boldsymbol{\sigma}}(\mathbf{x} - i\boldsymbol{\sigma}\mathbf{Y}). \quad (3.16)$$

Effectively, the full configuration on the complexified manifold is described by a real and positive function $Q(x_1, \sigma_1, \dots, x_n, \sigma_n)$. Each degree of freedom is augmented with an additional bit.⁸

Our proposal is to share each Fourier mode \mathbf{k} among 2^{m+1} branches, where the value of m and the concrete branches depend on the mode. For any such branch $\boldsymbol{\sigma}$, $Q_{\boldsymbol{\sigma}}$ is given by Q_+ in Eq. (3.13) with $\mathbf{Y} = Y\boldsymbol{\sigma}$ and an additional factor $1/2^m$. The concrete assignation of branches is as follows.

a) For a Fourier mode $\mathbf{k} = (k_1, \dots, k_n)$ with all k_i different from zero, only two branches are involved ($m = 0$) and Eq. (3.13) applies. One of the branches is that with $\sigma_i = \text{sign}(k_i)$, or equivalently, $k_i Y_i > 0$ for each i . The other branch is the opposite one, with all $k_i Y_i < 0$. This assignation of branches certainly guarantees that $\mathbf{k} \cdot \mathbf{Y}$ is never zero and complies with the uniformity criterion.

b) For Fourier modes in which some (but not all) of the k_i are zero: for the subset of k_i which are not zero the rule for the assignation of branch is as above (i.e., all $\sigma_i = \text{sign}(k_i)$)

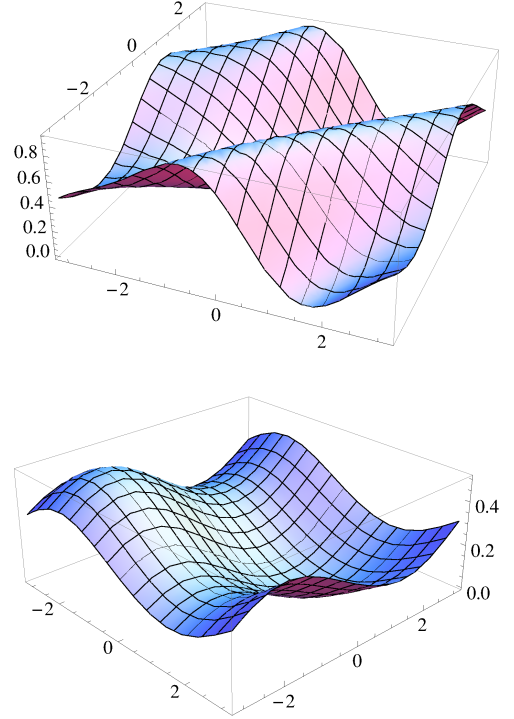


FIG. 1: Representations of $(1 + \beta \cos(x_1))(1 + \beta \cos(x_2))(1 + \beta \cos(x_1 - x_2))$ for $\beta = i$. (a) Exactly two sheets with $Y_2/Y_1 = \sqrt{2}$ and $Y_2 = 3.87$. The function $Q_+(\mathbf{x})$ is displayed. (b) Four sheets with $Y_1, Y_2 = \pm Y$ and $Y = 1.51$. The function $Q_{++}(\mathbf{x})$ is displayed.

or all $\sigma_i = -\text{sign}(k_i)$. For the vanishing k_i there is an ambiguity (completely analogous to the ambiguity in the choice of N_{\pm}). The most symmetric prescription is to assign half of the strength to each of the two possibilities $\sigma_i = \pm 1$. So a Fourier mode in which k_i vanishes for m values of i will be distributed among 2^{m+1} branches. Correspondingly Q_+ in Eq. (3.13) picks up a factor $1/2^m$.

c) The constant mode, $\mathbf{k} = 0$, is equally distributed among the 2^n branches, that is $N_{\boldsymbol{\sigma}} = 1/2^n$.⁹

Equivalently, for all $\boldsymbol{\sigma}$ and \mathbf{k} , Q_+ in Eq. (3.13) applies (with $\mathbf{Y} = Y\boldsymbol{\sigma}$) but with an additional factor. The factor is $1/2^m$ if m values $\sigma_i k_i$ vanish while the other are all positive or all negative. Otherwise the factor is zero.

As illustration, consider the two-dimensional distribution

$$P(x_1, x_2) \propto (1 + \beta \cos(x_1))(1 + \beta \cos(x_2))(1 + \beta \cos(x_1 - x_2)), \quad (3.17)$$

with $\beta = i$. This distribution admits a positive representation using exactly two sheets with an asymmetric choice $\mathbf{Y} \propto (1, \sqrt{2})$. The relation $Q_-(\mathbf{x}) = Q_+(-\mathbf{x})$ holds automati-

⁷ Such possibility is noted in [27] and it was also present in [22] where each Fourier mode is treated independently.

⁸ In counting degrees of freedom, this would be equivalent to duplicating the original coordinate range by joining two copies of it, for each coordinate. For instance, $[0, 2\pi] \rightarrow [0, 4\pi]$, or $\mathbb{R}^+ \rightarrow \mathbb{R}$. Unfortunately this picture does not work topologically, as the copies, say $[0, 2\pi]$ and $[2\pi, 4\pi]$, would not be related through any continuity condition.

⁹ Any other distribution with non negative $N_{\boldsymbol{\sigma}}$ would be valid, perhaps allowing a smaller \mathbf{Y} . The one proposed here is just the simplest one, and this also true for the prescription adopted in the case b).

cally. The optimal width, that is, such that $\min_x Q_+(x) = 0$, is obtained as $Y = (2.74, 3.87)$. The branch $Q_+(x)$ is displayed in Fig. 1a.

The alternative construction with four sheets, $Q_{\pm\pm}(x)$, attains a positive representation with $Y = (\pm 1.51, \pm 1.51)$, which having a smaller width represents an improvement over the previous asymmetric construction. Symmetry under $(x_1, x_2) \rightarrow (x_2, x_1)$ is automatic, and also $Q_{-\mp}(x) = Q_{+\pm}(-x)$ is fulfilled. The branch $Q_{++}(x)$ is displayed in Fig. 1b, the branch $Q_{+-}(x)$ has a similar shape, up to a reflection.

C. Representations from convolutions

The representations just described can be written as convolutions. Let us consider first the simple case in which problems coming from $k \cdot Y = 0$ can be neglected. The zero mode is treated separately as this singular term is always present. Straightforward reconstruction of the Fourier sum using the components in Eq. (3.13) gives

$$Q_+(x) = N_+ + 2\text{Re} \sum_{k \neq 0} \frac{e^{ik \cdot x} e^{k \cdot Y}}{2 \sinh(2k \cdot Y)} \int \frac{d^n x'}{(2\pi)^n} e^{-ik \cdot x'} P(x'). \quad (3.18)$$

In order to proceed, let us introduce the following function

$$\chi(\Omega) \equiv \frac{\Omega}{\Omega^2 - \Omega^{-2}} \quad (3.19)$$

and also the distribution

$$C(x; Y) \equiv \sum_{k \neq 0} e^{ik \cdot x} \chi(e^{k \cdot Y}). \quad (3.20)$$

This allows to express $Q_{\pm}(x)$ as convolutions:

$$\begin{aligned} Q_{\pm}(x) &= N_{\pm} + 2\text{Re}(C(x; \pm Y) * P(x)) \\ &= N_{\pm} + 2\text{Re}(C(\pm x; Y) * P(x)). \end{aligned} \quad (3.21)$$

(For convenience we denote $f(x) * g(x)$ what is usually denoted $(f * g)(x)$.) As is readily verified, the identities

$$\begin{aligned} \chi(\Omega)\Omega + \chi(\Omega^{-1})\Omega^{-1} &= 1, \\ \chi(\Omega)\Omega^{-1} + \chi(\Omega^{-1})\Omega &= 0, \end{aligned} \quad (3.22)$$

guarantee the fulfillment of Eq. (3.12). It should be noted that the expression using *real part* in Eq. (3.21) refers only to *real* x . Of course the analytical extension implied in Eq. (3.12) has to be applied *after* the real part is expanded in Eq. (3.21) as a linear combination of $P(x)$ and $P^*(x)$.

We can turn now to the improved construction using 2^n branches. Again the zero mode is treated separately, only subject to the conditions

$$1 = \sum_{\sigma} N_{\sigma}, \quad N_{\sigma} \geq 0. \quad (3.23)$$

For the remaining Fourier modes the expression in Eq. (3.18) still holds with $Y = Y\sigma$ and taking into account that not all modes contribute to each branch σ : In principle, a given

mode $k = (k_1, \dots, k_n)$ contributes only to the branch with all σ_i equal to $\text{sign}(k_i)$ or all opposite. When some k_i are zero, these are equally distributed between the $\sigma_i = 1$ and $\sigma_i = -1$ options.

In this way, the functions $Q_{\sigma}(x)$ can be written as convolutions in the form

$$Q_{\sigma}(x) = N_{\sigma} + 2\text{Re}(C(\sigma * x; Y) * P(x)), \quad \sigma = (\pm, \dots, \pm), \quad (3.24)$$

where we have defined

$$\sigma * x \equiv (\sigma_1 x^1, \dots, \sigma_n x^n), \quad (3.25)$$

and

$$C(x; Y) \equiv \sum_k e^{ik \cdot x} \chi(e^{Y \sum_i k_i}) \Theta(k). \quad (3.26)$$

The function $\Theta(k)$ selects the Fourier modes contributing to the branch $(+, \dots, +)$,¹⁰

$$\Theta(k) \equiv \begin{cases} \prod_{i=1}^n \theta(k_i) + \prod_{i=1}^n \theta(-k_i) & (k \neq 0) \\ 0 & (k = 0) \end{cases}, \quad (3.27)$$

$\theta(x)$ being the Heaviside step function with $\theta(0) = 1/2$.

D. Complex representations and linearity

Through Eq. (2.2), any (in general complex) distribution ρ on the complexified manifold defines an associated complex probability P on the real manifold. (Of course, as for the observables, this assumes some class of sufficiently well behaved distributions ρ .) Let us denote by K the corresponding projection operator, that is,

$$K\rho = P. \quad (3.28)$$

K is the transposed operator of that carrying out the analytical extension of an observable from the real to the complex manifold; with obvious notation for the bilinear scalar product:

$$\langle A(x) \rangle_P = \langle A, P \rangle = \langle A, K\rho \rangle = \langle K^T A, \rho \rangle = \langle A(z) \rangle_{\rho}. \quad (3.29)$$

For sufficiently well behaved distributions ρ on \mathbb{C}^n K can be expressed as [14, 32]

$$P(x) = \int d^n y e^{-iy \cdot \nabla_x} \rho(x, y) \equiv \int d^n y \rho(x - iy, y). \quad (3.30)$$

(This is a straightforward consequence of $\langle A(x), P(x) \rangle = \langle A(x + iy), \rho(z) \rangle$.) Eq. (3.12) illustrates this relation when $\rho(z)$ is that in Eq. (3.11).

Obviously K is a linear operator, so if some ρ_i are representations of some P_i ,

$$K\rho_i = P_i, \quad i = 1, \dots, N, \quad (3.31)$$

¹⁰ The function $\Theta(\sigma * k)$ does the same job for the branch σ .

automatically a complex linear combination

$$P = \sum_i a_i P_i \quad (3.32)$$

admits the corresponding combination as representation

$$\rho = \sum_i a_i \rho_i, \quad P = K\rho. \quad (3.33)$$

Unfortunately, even if all the ρ_i are real, such ρ will be complex in general since the a_i are complex.

Abstracting what has been implicitly done in the previous subsections, one can proceed as follows. The constant mode is treated separately and added a posteriori. So we consider here complex distributions with *zero normalization*: $P = \sum_i a_i P_i$ where the P_i , and hence P , integrate to zero.

Let us introduce the anti-analytic version of K , which will be denoted by \bar{K} and is also linear, through the relation

$$\bar{K}\rho = (K\rho^*)^*. \quad (3.34)$$

Now given a basis (actually any linear set) $P_i(x)$ we associate a set of *complex* representations $\hat{P}_i(z)$ subject to the (linear) requirements

$$P_i = K\hat{P}_i, \quad 0 = \bar{K}\hat{P}_i. \quad (3.35)$$

That is, the analytic projections of the \hat{P}_i yield P_i (i.e., the \hat{P}_i are proper representation of P_i albeit complex) while their anti-analytic projections vanish. Then, obviously

$$\hat{P} = \sum_i a_i \hat{P}_i \quad (3.36)$$

is also a (complex) representation of P , $P = K\hat{P}$. Note that the second equation in (3.35) is equivalent to

$$0 = K\hat{P}_i^*. \quad (3.37)$$

Hence $0 = K\hat{P}^*$, and

$$\rho = 2\text{Re}(\hat{P}) \quad (3.38)$$

is, by construction, a *real* representation of P ,

$$P = K\rho. \quad (3.39)$$

(Recall that constant modes still have to be added at both sides of Eq. (3.38) to have properly normalized distributions.)

The Eqs. (3.35) admit many solutions for given P_i . For instance the (symmetric) two-branch construction depends on the width parameter Y , and many more solutions exists. It is interesting that unlike ρ , the complex representations \hat{P} or \hat{P}_i preserve information on the phases of P and P_i , respectively. Another remark is that if P_i has some symmetry, one can impose the same symmetry on its complex representation \hat{P}_i , so each symmetry type (irreducible representation of the symmetry group) can be represented independently, thanks to the linearity of the construction.

Another comment refers to this construction in the noncompact case. The expression in Eq. (3.6) holds equally well for a

normalized complex probability $P(x)$ defined on \mathbb{R} , using the Fourier components $\tilde{P}(k)$ there,

$$P(x) = \int \frac{dk}{2\pi} e^{ikx} \tilde{P}(k), \quad \int dx P(x) = 1. \quad (3.40)$$

The $k \rightarrow 0$ limits of $\tilde{Q}_{\pm}(k)$ in Eq. (3.6) exist, since $\tilde{P}(0)$ is a real number. As a consequence $\tilde{Q}_{\pm}(0)$ take well defined values, rather than being free parameters as in the compact case.

The functions $Q_{\pm}(x)$ receive (linear) contributions from $P(x)$ and $P^*(x)$, and we can denote $\tilde{Q}_{\pm}(x)$ the component coming only from P (analogous to \hat{P} , as compared to $\rho = 2\text{Re}(\hat{P})$). In this case one finds that the Fourier modes

$$\hat{\tilde{Q}}_{\pm}(k) = \pm \frac{e^{\pm kY}}{2 \sinh(2kY)} \tilde{P}(k), \quad (3.41)$$

display a pole at $k = 0$. This implies that the complex representations $\tilde{Q}_{\pm}(x)$ are not convergent at infinity. More precisely, their real parts, $Q_{\pm}(x)$, are convergent but their imaginary parts are not.

In general, in the noncompact case, complex representations \hat{P}_i corresponding to *normalized* P_i , will produce complex combinations $\sum_i a_i \hat{P}_i$ which will not be properly convergent, however, the divergence cancels in their real parts provided the normalization $\sum_i a_i$ is a real number.

Let us note that the infrared divergence must necessarily be present in \hat{P} (this is clear in Eq. (3.41), since $\tilde{P}(0) = 1$). This comes from a conflict in Eq. (3.35) in the noncompact case. In the compact case, the constant mode was cleanly separated and all distributions in Eq. (3.35) were assumed to have zero normalization. The same cannot be done in the noncompact case. If the constant mode cannot be extracted one finds an incompatibility in Eq. (3.35). To see this let us denote by \hat{P}_0 and \hat{P}_{00} the operators yielding the normalization of distributions on \mathcal{M} and \mathcal{M}^c , respectively ($\hat{P}_0 = \int dx$ and $\hat{P}_{00} = \int d^2z$ for $\mathcal{M} = \mathbb{R}$). These operators fulfill the identities

$$\hat{P}_0 K = \hat{P}_0 \bar{K} = \hat{P}_{00}. \quad (3.42)$$

Applying them to

$$P = K\hat{P}, \quad 0 = \bar{K}\hat{P}. \quad (3.43)$$

one finds

$$1 = \hat{P}_0 P = \hat{P}_{00} \hat{P}, \quad 0 = \hat{P}_{00} \hat{P}. \quad (3.44)$$

The conflict results in a singularity in the imaginary part of \hat{P} at the constant mode.

IV. LOCALIZED REPRESENTATIONS ON LIE GROUPS

In this section we aim at extending the previous constructions to non necessarily Abelian Lie groups. Eventually we will limit our study to compact groups because too general (group) representations of noncompact groups would be intractable. Nevertheless, it can be conjectured that our results apply also to a complex probability P defined on any

Lie group G , provided P is spanned by a set of well behaved representations of G (e.g., bounded representations). The case $G = (\mathbb{R}^n, +)$ and $P(x)$ admitting a Fourier decomposition in terms of $e^{ik \cdot x}$, for $k \in \mathbb{R}^n$ (as opposed to $k \in \mathbb{C}^n$) is such an example.

A. Representations on groups

For definiteness we will assume a matrix group,

$$G = \{g(\mathbf{a}) = e^{\mathbf{a} \cdot \mathbf{T}}, \mathbf{a} \in \mathbb{R}^n\}, \quad (4.1)$$

where the matrices T_i ($i = 1, \dots, n$) are the group generators and a^i ($i = 1, \dots, n$) are the normal coordinates of the element g . New admissible real coordinate systems are derived by means of *real analytic* changes of variables.

The complexified group G^c is obtained by taking complex values for the coordinates,

$$G^c = \{g(\mathbf{a}) = e^{\mathbf{a} \cdot \mathbf{T}}, \mathbf{a} \in \mathbb{C}^n\}. \quad (4.2)$$

The analytically extended observables are defined on G^c through analytical extension with respect to their dependence on the coordinates. (The extension does not depend on the concrete coordinates used as long as they belong to the class of admissible ones.)

Given a positive measure $d\mu(g)$ on G , one can define complex distributions $P(g)$ on G and corresponding expectation values. The factor between two different choices of measure can be reabsorbed in the distribution, so without loss of generality, we will use the right-invariant Haar measure of G . For compact G we adopt the normalized measure

$$\int_G dg = 1 \quad (\text{compact } G). \quad (4.3)$$

Likewise, we take the right-invariant measure on G^c . The complexified group is never compact, but will be unimodular if G is.¹¹ The concept of representation works as before, as dictated by Eq. (2.2).

We will need to introduce the (complex) conjugate element \bar{g} of a given $g \in G^c$. This is defined by

$$g = g(\mathbf{a}), \quad \bar{g} = g(\mathbf{a}^*) \quad \mathbf{a} \in \mathbb{C}^n. \quad (4.4)$$

This conjugation is a group automorphism in G^c and its definition does not depend on the particular coordinates used in G . Also note that \bar{g} needs not coincide with g^* (the conjugate matrix in a matrix group) unless $\mathbf{T}^* = \mathbf{T}$.

An important property of the conjugation is that, for any (group) representation $D(g)$ of G and $D^*(g) = D(g)^*$ its conjugate representation, upon analytical extension into G^c ,

$$(D(g))^* = D^*(\bar{g}) \quad g \in G^c. \quad (4.5)$$

Obviously, the set of autoconjugated (real) elements is G itself,

$$g = \bar{g} \quad \text{iff} \quad g \in G. \quad (4.6)$$

The subset of *purely imaginary* elements of G^c , which we denote G_I , can be naturally defined as

$$g \in G_I \quad \text{iff} \quad \bar{g} = g^{-1}. \quad (4.7)$$

In normal coordinates G_I are those elements of G^c with purely imaginary coordinates. In the non Abelian case G_I is not a subgroup of G^c , however if $g \in G$, $gG_Ig^{-1} = G_I$. Also, if $h \in G_I$, $h^s \in G_I$, for $s \in \mathbb{R}$.¹²

B. Two-branch representations

We will not need very general distributions on G^c , rather we use a two-branch approach (with suitable variations in the higher dimensional case, as in Sec. III B 2). That is, for a given (normalized) complex probability $P(g)$

$$1 = \int_G dg P(g), \quad (4.8)$$

we seek two positive distributions $Q_{\pm}(g)$ on G in such a way that they define a representation of $P(g)$, by means of the relation

$$P(g) = Q_+(gg_+) + Q_-(gg_-) \quad \forall g \in G, \quad (4.9)$$

where $g_{\pm} \in G^c$ are two parameters of the construction, and $Q_{\pm}(gg_{\pm})$ refer to the analytical extension of $Q_{\pm}(g)$ into the complexified group. Indeed, using the right-invariance of the measure,

$$\begin{aligned} \langle A \rangle_P &= \int_G dg A(g) P(g) = \int_G dg A(g) \sum_{\sigma=\pm} Q_{\sigma}(gg_{\sigma}) \\ &= \int_G dg \sum_{\sigma=\pm} Q_{\sigma}(g) A(gg_{\sigma}^{-1}) \\ &= N_+ \langle A(gg_+^{-1}) \rangle_{Q_+} + N_- \langle A(gg_-^{-1}) \rangle_{Q_-}, \end{aligned} \quad (4.10)$$

where N_{\pm} denote the normalizations of Q_{\pm} ,

$$N_{\pm} = \int_G dg Q_{\pm}(g), \quad (4.11)$$

with

$$1 = N_+ + N_-, \quad N_{\pm} \geq 0. \quad (4.12)$$

Eq. (4.10) implies that the expectation value of A can be obtained by importance sampling of the two positive distributions $Q_{\pm}(g)$ defined on G . The representation $\rho(g)$ itself has

¹¹ Since the invariant measure on G^c is $|\sigma(\mathbf{a})|^2 d^n a d^n a^*$ when the invariant measure on G is $\sigma(\mathbf{a}) d^n a$.

¹² For $G = \text{SU}(2)$, the rotation group, $G^c = \text{SL}(2, \mathbb{C})$ is the Lorentz group and G_I is the set of boosts.

support on two copies of G contained in G^c , namely, Gg_+^{-1} and Gg_-^{-1} . Therefore the elements g_\pm represent the displacements away from G into G^c .

In Eq. (4.9) we have arbitrarily chosen the shift to act on the right. Of course everything would be analogous with $Q_+(g_+g) + Q_-(g_-g)$. Also possible would be (for a unimodular group)

$$P(g) = Q_+(g'_+gg_+) + Q_-(g'_-gg_-). \quad (4.13)$$

We do not explore this latter possibility as it is technically more complicated with no obvious advantage.

It is clear that there is no solution to Eq. (4.9) (with positive Q_\pm) if $g_\pm \in G$, unless P is already a positive distribution. As discussed in Sec. II C, the representation $\rho(g)$ must have some support sufficiently far from the real manifold (the group G in this case); a minimal width is required for any positive representation ρ .

The complex distribution P is equivalent (has the same information as) to two real functions, so it can be expected that for given g_\pm , the two real functions Q_\pm are essentially unique. To actually determine the two branches Q_\pm we apply the approach developed in Sec. III D as follows.

The (group) representations of a group span the space of complex functions defined on that group (i.e., its regular representation). So general distributions $P(g)$ can be expanded as linear combinations of (group) representations $D^R(g)$ of G , i.e., $P(g) \sim \sum_R P^R D^R(g)$.

In order to cleanly separate the normalization mode (constant mode) in P , we will assume in what follows that G is a *compact group*, hence our complex normalized probability $P(g)$ can be expressed as

$$\begin{aligned} P(g) &= 1 + \sum_{R \neq 1} \sum_{\alpha, \beta} (P^R)^\beta_\alpha D^R(g)^\alpha_\beta \\ &= 1 + \sum_{R \neq 1} \text{tr}(P^R D^R(g)). \end{aligned} \quad (4.14)$$

The P^R are constant complex matrices of the same dimension as the representation R . We have separated the trivial (or singlet) representation $D^{R=1}(g) \equiv 1$ which must carry weight 1 if P is normalized.

As follows from the Peter-Weyl theorem, the set of irreducible representations (irreps) form an orthonormal basis for the regular representation and we could take the R to be irreducible, however, such assumption is not strictly needed for our construction, so we will only assume that R does not contain the trivial representations in its decomposition into irreps, therefore

$$\int_G dg D^R(g) = 0 \quad (R \neq 1). \quad (4.15)$$

To apply the scheme of Sec. III D, we will seek *complex* representations for each component R in P , fulfilling the conditions in Eqs. (3.35). That is, for each $R \neq 1$ we seek two functions $\hat{Q}_\pm^R(g)$ of the form

$$\hat{Q}_\pm^R(g) = \text{tr}(Q_\pm^R D^R(g)), \quad (4.16)$$

where Q_\pm^R are two matrices to be determined. Then the real distributions

$$Q_\pm^R(g) = 2\text{Re} \hat{Q}_\pm^R(g) \quad g \in G \quad (4.17)$$

are the two real branches in the representation of the component R of $P(g)$ and

$$Q_\pm(g) = N_\pm + \sum_{R \neq 1} Q_\pm^R(g). \quad (4.18)$$

The two functions $\hat{Q}_\pm^R(g)$ are determined through Eq. (3.35).¹³ The representation condition on $\hat{Q}_\pm^R(g)$ (first relation in Eq. (3.35)) becomes (using Eq. (4.16))

$$\text{tr}(P^R D^R(g)) = \sum_{\sigma=\pm} \hat{Q}_\sigma^R(gg\sigma) = \sum_{\sigma=\pm} \text{tr}(Q_\sigma^R D^R(g) D^R(g\sigma)), \quad (4.19)$$

that is

$$P^R = D^R(g_+) Q_+^R + D^R(g_-) Q_-^R. \quad (4.20)$$

To impose the second relation in Eq. (3.35), note that

$$\hat{Q}_\pm^R(g)^* = \text{tr}(Q_\pm^{R*} D^{R*}(g)), \quad g \in G, \quad (4.21)$$

where D^{R*} is the conjugate representation of D^R . Then Eq. (3.37) takes the form

$$0 = \sum_{\sigma=\pm} \text{tr}(Q_\sigma^{R*} D^{R*}(g) D^{R*}(g\sigma)). \quad (4.22)$$

Taking complex conjugation and using Eq. (4.5) yields

$$0 = \sum_{\sigma=\pm} \text{tr}(Q_\sigma^R D^R(g) D^R(\bar{g}\sigma)), \quad (4.23)$$

which provides a second equation on Q_\pm^R :

$$0 = D^R(\bar{g}_+) Q_+^R + D^R(\bar{g}_-) Q_-^R. \quad (4.24)$$

Assuming that the required matrices are invertible, the system of Eqs. (4.20) and (4.24) can be solved to give

$$Q_\pm^R = (D^R(g_\mp^{-1} g_\pm) - D^R(\bar{g}_\mp^{-1} \bar{g}_\pm))^{-1} D^R(g_\mp^{-1}) P^R. \quad (4.25)$$

Equivalently,

$$Q_\pm^R = (1 - D^R(g_\pm^{-1} g_\mp \bar{g}_\mp^{-1} \bar{g}_\pm))^{-1} D^R(g_\pm^{-1}) P^R. \quad (4.26)$$

So a solution is obtained whenever the matrix $D^R(g_\pm^{-1} g_\mp \bar{g}_\mp^{-1} \bar{g}_\pm)$ has no eigenvalue $\lambda = 1$. If it has, there can still be solutions if P has no component along those eigenvectors. We come back to this crucial question in Sec. V. For the time being we will assume that the required matrices are indeed invertible. As always the trivial

¹³ Noting that Eq. (4.9) expresses $P = K\rho$.

representation (constant mode) has been explicitly extracted (since certainly all eigenvalues $\lambda = 1$ when $R = 1$).

Since those components of g_{\pm} along G are ineffective, the most efficient choice, in principle, corresponds to taking purely imaginary displacements. Hereafter we adopt this prescription, $g_{\pm} \in G_I$, and also choose a symmetric disposition of the two shifts, $g_+ = g_-^{-1}$:

$$h \equiv g_+ = g_-^{-1} = \bar{g}_+^{-1} = \bar{g}_- \in G_I. \quad (4.27)$$

Then Eq. (4.9) becomes

$$P(g) = Q_+(gh) + Q_-(gh^{-1}) \quad \forall g \in G, \quad (4.28)$$

and

$$\langle A \rangle_P = N_+ \langle A(gh^{-1}) \rangle_{Q_+} + N_- \langle A(gh) \rangle_{Q_-}. \quad (4.29)$$

Also Eqs. (4.20) and (4.24) become

$$\begin{aligned} P^R &= D^R(h) Q_+^R + D^R(h^{-1}) Q_-^R \\ 0 &= D^R(h^{-1}) Q_+^R + D^R(h) Q_-^R. \end{aligned} \quad (4.30)$$

In addition Eq. (4.25) becomes

$$Q_+^R = \chi(D^R(h)) P^R, \quad Q_-^R = \chi(D^R(h^{-1})) P^R, \quad (4.31)$$

where χ is the function introduced in Eq. (3.19) and $\chi(D^R(h))$ is a matrix of the same dimension as R . Therefore, the two branches for the representation of $P(g)$ can be compactly written as

$$Q_{\pm}(g) = N_{\pm} + 2\text{Re} \sum_{R \neq 1} \text{tr}[P^R D^R(g) \chi(D^R(h^{\pm 1}))]. \quad (4.32)$$

Because G is compact and its representations R are unitary, the matrices $D^R(g)$ are unitary, while $D^R(h)$ (and hence $\chi(D^R(h))$) are hermitian. This follows from the identity

$$D^R(g^{-1}) = D^R(g)^{-1} = D^R(g)^{\dagger} \quad \forall g \in G^c \quad (R \text{ unitary}). \quad (4.33)$$

Once again, for sufficiently large h (assuming no $\lambda = 1$ eigenvalues are involved) $\chi(D^R(h))$ goes to zero and only the singlet (trivial representation) mode remains in Eq. (4.32), implying that eventually Q_{\pm} become non negative.

Of course the case $G = \text{U}(1)^{\times n}$ studied in Sec. III conforms to this general scheme: the normal coordinates are $\mathbf{a} = \mathbf{x}$ in G and $\mathbf{a} = \mathbf{z}$ in $G^c = (\text{U}(1) \times \mathbb{R})^{\times n}$. Also, $R = \mathbf{k}$, $D^R(g) = e^{i\mathbf{k} \cdot \mathbf{x}}$ and $P^R = \tilde{P}_{\mathbf{k}}$. Furthermore, h has coordinates $-i\mathbf{Y}$ and so $D^R(h) = e^{\mathbf{k} \cdot \mathbf{Y}}$. In this way Eq. (4.32) reproduces Eq. (3.18).

C. An SU(2) example

Let us consider the following complex probability on $G = \text{SU}(2)$

$$P(g) = 1 + \text{tr}(pg) \quad g \in \text{SU}(2). \quad (4.34)$$

Here p is a constant complex 2×2 matrix. Letting $h \in \text{SU}(2)_I$, a direct application of the previous results gives

$$Q_{\pm}(g) = N_{\pm} \pm 2\text{Re tr}(h^{\pm 1}(h^2 - h^{-2})^{-1}pg). \quad (4.35)$$

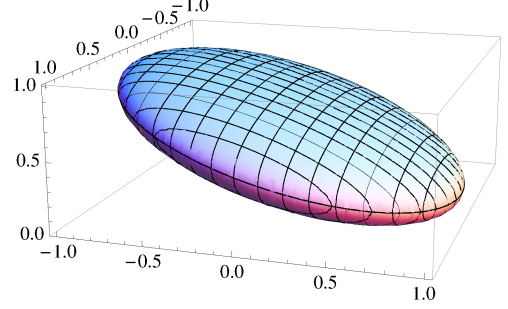


FIG. 2: For $g \in \text{SU}(2)$, function $Q_+(g)$ on the plane (a_1, a_3) with $a_2 = 0$, for $P(g) = 1 + \beta \text{tr}(g)$ with $\beta = 1 - i$, and $h = \text{diag}(e^Y, e^{-Y})$ with $Y = 3.5$, and $N_+ = 1/2$.

To be more explicit, let

$$\begin{aligned} g &= \cos(\psi/2) - i \sin(\psi/2) \hat{\psi} \cdot \sigma, \\ p &= p_0 + \mathbf{p} \cdot \sigma, \\ h &= \cosh(Y) + \sinh(Y) \hat{\mathbf{Y}} \cdot \sigma, \end{aligned} \quad (4.36)$$

where p_0 and \mathbf{p} can be complex and ψ and \mathbf{Y} are real. Then

$$Q_{\pm}(g) = A_{\pm} a_0 + \mathbf{B}_{\pm} \cdot \mathbf{a} \quad (4.37)$$

with

$$a_0 = \cos(\psi/2), \quad \mathbf{a} = \sin(\psi/2) \hat{\psi} \quad (4.38)$$

and

$$\begin{aligned} A_{\pm} &= \frac{\text{Re}(p_0)}{\cosh(Y)} \pm \frac{\hat{\mathbf{Y}} \cdot \text{Re}(\mathbf{p})}{\sinh(Y)} \\ \mathbf{B}_{\pm} &= \pm \frac{\hat{\mathbf{Y}} \text{Im}(p_0)}{\sinh(Y)} \pm \frac{\hat{\mathbf{Y}} \times \text{Re}(\mathbf{p})}{\sinh(Y)} + \frac{\text{Im}(\mathbf{p})}{\cosh(Y)}. \end{aligned} \quad (4.39)$$

As an illustration, in Fig. 2 we show the function $Q_+(g)$ for

$$P(g) = 1 + \beta \text{tr}(g), \quad \beta = 1 - i, \quad (4.40)$$

using $\hat{\mathbf{Y}} = (0, 0, 1)$ and $Y = 3.5$, and $N_+ = \frac{1}{2}$. $\text{SU}(2)$ is a three-sphere, $a_0^2 + \mathbf{a}^2 = 1$, so Q_+ as a function of \mathbf{a} is two-valued. The plot displays $Q_+(a_1, 0, a_3)$, the submanifold $a_2 = 0$ being a two-sphere.

It is interesting to note that in any $\text{U}(N)$ group the complex probabilities of the type in Eq. (4.34) can be reduced to a standard form before representation. The matrix p can be written as

$$p = u_L d u_R, \quad u_{L,R} \in \text{U}(N), \quad d \text{ diagonal and non negative}, \quad (4.41)$$

so that

$$P(g) = 1 + \text{tr}(d u_R g u_L). \quad (4.42)$$

Then it is sufficient to find representations $\rho'(g)$ for

$$P'(g) = 1 + \text{tr}(dg), \quad (4.43)$$

and afterwards undo the left and right translations

$$\rho(g) = \rho'(u_R g u_L). \quad (4.44)$$

In the case of $\text{SU}(N)$,

$$p = e^{i\varphi} u_L d u_R, \quad u_{L,R} \in \text{SU}(N), \quad d \text{ diagonal and non negative.} \quad (4.45)$$

(φ real.) In particular for $\text{SU}(2)$ the most general case required is $p = e^{i\varphi} a(1 + \cos \theta \sigma_3)$, $a > 0$, $\theta, \varphi \in \mathbb{R}$.

D. Representations through convolutions

The functions $Q_{\pm}(g)$ can also be obtained from convolution of $P(g)$ with a fixed kernel. To do this, we express P in terms of irreducible group representations, μ , as

$$P(g) = \sum_{\mu} \text{tr}(P^{\mu} D^{\mu}(g)), \quad (4.46)$$

$$P^{\mu} = n_{\mu} \int_G dg P(g) D^{\mu}(g^{-1}),$$

where n_{μ} denotes the dimension of the irrep μ . Using the expression of P^{μ} to work out Eq. (4.32), one obtains¹⁴

$$Q_{\pm}(g) = N_{\pm} + 2\text{Re} \int dg' P(g') C(g'^{-1}g; h^{\pm 1}) \quad (4.47)$$

$$= N_{\pm} + 2\text{Re}(P(g) * C(g; h^{\pm 1})),$$

with

$$C(g; h) = \sum_{\mu \neq 1} n_{\mu} \text{tr}[D^{\mu}(g) \chi(D^{\mu}(h))]. \quad (4.48)$$

Eqs. (4.47) and (4.48) generalize Eqs. (3.21) and (3.20), respectively.

Summations on μ or μ^* (the conjugate irrep) are equivalent within the trace in Eq. (4.48). Using this rearrangement, along with Eq. (4.33) and

$$D^{R*}(g) = D^R(g^{-1})^T \quad \forall g \in G^c \quad (R \text{ unitary}) \quad (4.49)$$

one can easily establish the following identities

$$C(g; h^{-1}) = C(g^{-1}; h) = C(g; h)^*. \quad (4.50)$$

¹⁴ The group convolution

$$(A * B)(g) \equiv \int_G dg' A(g') B(g'^{-1}g) = \int_G dg' A(gg'^{-1}) B(g')$$

is not commutative in general.

E. Representations in matrix groups

Let $G \subseteq \text{U}(N)$, and g^i_j the matrix elements of $g \in G$ ($i, j = 1, \dots, N$). The (group) representations of G can be obtained from tensor product of the basic representations g and g^* . (Note that such product representations will be reducible in general.)

In the simplest case in which only g is involved

$$P(g) = \sum_{n=0}^{\infty} p^{j_1 \dots j_n}_{i_1 \dots i_n} g^{i_1}_{j_1} \dots g^{i_n}_{j_n}, \quad (4.51)$$

where $p^{j_1 \dots j_n}_{i_1 \dots i_n}$ are complex coefficients. This is a decomposition of P into group representations of the type $D^R(g) = g \otimes \dots \otimes g$ (n factors),

$$D^R(g)^{i_1 \dots i_n}_{j_1 \dots j_n} = g^{i_1}_{j_1} \dots g^{i_n}_{j_n} \equiv (g^{\otimes n})^{i_1 \dots i_n}_{j_1 \dots j_n}, \quad (4.52)$$

and Eq. (4.32) applies

$$Q_+(g) = N_+ + 2\text{Re} \sum_{n=1}^{\infty} \hat{Q}_n(g), \quad (4.53)$$

with

$$\hat{Q}_n(g) = p^{j_1 \dots j_n}_{i_1 \dots i_n} (g^{\otimes n})^{i_1 \dots i_n}_{k_1 \dots k_n} \chi(h^{\otimes n})^{k_1 \dots k_n}_{j_1 \dots j_n}. \quad (4.54)$$

The contribution to $Q_-(g)$ is analogous, using h^{-1} instead of h . Also note that because R is unitary, h is hermitian.

Let us assume that $h \in G_I$ is a diagonal matrix,

$$h = \text{diag}(\omega_1, \dots, \omega_N). \quad (4.55)$$

The ω_i are real (and moreover positive for a connected group). In this case $h^{\otimes n}$ and $\chi(h^{\otimes n})$ are also diagonal and $\hat{Q}_n(g)$ takes a simple form

$$\hat{Q}_n(g) = p^{j_1 \dots j_n}_{i_1 \dots i_n} g^{i_1}_{j_1} \dots g^{i_n}_{j_n} \chi(\omega_{j_1} \dots \omega_{j_n}). \quad (4.56)$$

Denoting by Ω the argument of the function χ , $\Omega = \omega_{j_1} \dots \omega_{j_n}$, we can see that it picks up a factor ω_j for each factor g^i_j in the representation R .

More generally, $R = g^{\otimes n} \otimes g^{*\otimes m}$. The corresponding right translation with $h \in G_I$ is

$$g \mapsto gh \quad g^* = g^{-1T} \mapsto g^* h^{-1T}. \quad (4.57)$$

This implies that Ω picks up a factor ω_j for each factor g^i_j in R , and a factor ω_j^{-1} from each factor g^{*i}_j . That is, a term

$$P(g) = g^{i_1}_{j_1} \dots g^{i_n}_{j_n} g^{*l_1}_{k_1} \dots g^{*l_m}_{k_m}, \quad (4.58)$$

gives a contribution

$$\hat{Q}(g) = g^{i_1}_{j_1} \dots g^{i_n}_{j_n} g^{*l_1}_{k_1} \dots g^{*l_m}_{k_m} \chi(\omega_{j_1} \dots \omega_{j_n} \omega_{k_1}^{-1} \dots \omega_{k_m}^{-1}). \quad (4.59)$$

Similar formulas hold in more general cases.¹⁵ Also note that $g^{\otimes n}$ suffices for $SU(2)$ since g and $g^* = \sigma_y g \sigma_y$ are equivalent representations in this case.

Another observation is that Ω may be equal to 1 for some components and the previous formulas do not directly apply there. This will certainly happen when R contains the trivial representation in its reduction, but not only then. This problem is addressed in Sec. V.

If a configuration of the real manifold consists of n variables, (g^1, \dots, g^n) , each of them an element of the group $G_1 \subseteq U(N)$, the complex probability is defined on the group $G = G_1 \times \dots \times G_1$ (n factors) and $g = g^1 \dots g^n$. The formulas apply as before, and for instance, a term of the form

$$P(g) = g^{1i_1}_{j_1} g^{1*i_2}_{j_2} g^{2i_3}_{j_3}, \quad (4.60)$$

with diagonal h with parameters ω_j , would yield a contribution

$$\hat{Q}(g) = g^{1i_1}_{j_1} g^{1*i_2}_{j_2} g^{2i_3}_{j_3} \chi(\omega_{j_1} \omega_{j_2}^{-1} \omega_{j_3}). \quad (4.61)$$

It should be noted that a discussion similar to that in Sec. IIIB 2 can be (and should be) done here to restore uniformity with respect to the n variables, resulting in a total of 2^n branches, instead of 2. In the Abelian case, bifurcation of the variables solved the problem of singular terms ($\mathbf{k} \cdot \mathbf{Y} = 0$ denominators). A crucial difference with the Abelian case is that the presence of singular components (not invertible matrices in Eq. (4.32)) is not automatically solved by bifurcation in the nonabelian case, so we defer the discussion to Sec. V.

When the element $h \in G_I$ is not directly diagonal but it is diagonalizable within G ,¹⁶ a practical way to proceed is as follows. Let

$$h = U h_z U^{-1}, \quad U \in G, \quad h_z \in G_I \text{ and diagonal}, \quad (4.62)$$

and let

$$P'(g) \equiv P(U g U^{-1}). \quad (4.63)$$

Then

$$\hat{Q}(g) = \hat{Q}'(U^{-1} g U), \quad (4.64)$$

where $\hat{Q}'(g)$ is the complex representation associated to $P'(g)$, constructed using the diagonal h_z as described above. Indeed,

using Eq. (4.32),

$$\begin{aligned} \hat{Q}(g) &= \text{tr}(P^R D^R(g) \chi(D^R(h))) \\ &= \text{tr}(P^R D^R(g) D^R(U) \chi(D^R(h_z)) D^R(U)^{-1}) \\ &= \text{tr}(P'^R D^R(U^{-1} g U) \chi(D^R(h_z))) = \hat{Q}'(U^{-1} g U). \end{aligned} \quad (4.65)$$

V. REMOVAL OF SINGULAR KERNELS AND EXAMPLES

A. Singular kernels

The first expression in Eq. (4.31) can be rewritten as

$$Q_+^R = D^R(h)^3 (D^R(h)^4 - 1)^{-1} P^R \quad (5.1)$$

and similarly for Q_-^R with h^{-1} . Hence there is a proper solution when $D^R(h)$ has no $\lambda = 1$ eigenvalues¹⁷ or, if it has, P^R has no components along the corresponding eigenvectors. Otherwise we meet an obstruction to solving Eq. (4.30).

As already noted, when a probability $P(g)$ is complex, the support of any of its *real* representations must necessarily extend beyond G into the complexified manifold. In the two-branch approach the pushing into G^c is carried out by h (or more generally g_{\pm}). An obstruction arises when some components of P^R are not moved by $D^R(h)$ (unless they happen to be already positive). The obstruction takes place when some components of P remain invariant under the action of h , i.e., when h does not act *effectively* on all components of P . This is quite clear in the Abelian case $U(1)^{\times n}$ discussed in Sec. IIIB 1. There, an obstruction was met for Fourier modes such that $\mathbf{k} \cdot \mathbf{Y} = 0$. They correspond to the Fourier components $e^{i\mathbf{k} \cdot \mathbf{x}}$ of P which remain invariant under the imaginary translation $-i\mathbf{Y}$.

An important observation is that, in the nonabelian case, the obstruction cannot be removed by a clever choice of h (or even g_{\pm} outside G_I). To see this it suffices to consider the case $G = SU(2)$. If $R = j$ is a half-integer representation, $D^j(h)$ has no eigenvalue equal to 1, since the operator J_z has no zero eigenvalues, and the same is true of $J_n = \hat{n} \cdot \mathbf{J}$; so for those irreps any choice of rotation axis provides a solution.¹⁸ However, for integer j , J_n has exactly one zero eigenvalue. This means that no matter how the (complex) rotations are chosen $D^j(h)$ will have an eigenvalue equal to one for some eigenvector. We conclude that for integer j the obstruction cannot be avoided by just a better choice of the element h . For h imaginary the rotation angle is imaginary and the rotation axis \hat{n} is

¹⁵ In $GL(N, \mathbb{C})$, a non compact group, R would be obtained as a direct product of basic representations g, g^*, g^{-1T} and $g^{-1\dagger}$. The corresponding right translation with $h \in G_I$ (which is no longer hermitian) would be $gh, g^* h^{-1*}, g^{-1T} h^{-1T}$ and $g^{-1\dagger} h^{\dagger}$, respectively. So, for instance, a term of the form $P(g) = g^{i_1}_{j_1} g^{*i_2}_{j_2} (g^{-1T})^{i_3}_{j_3} (g^{-1\dagger})^{i_4}_{j_4}$ would produce a contribution $\hat{Q}(g) = g^{i_1}_{j_1} g^{*i_2}_{j_2} (g^{-1T})^{i_3}_{j_3} (g^{-1\dagger})^{i_4}_{j_4} \chi(\omega_{j_1} \omega_{j_2}^{-1*} \omega_{j_3}^{-1} \omega_{j_4}^*)$.

¹⁶ When $G \subset U(N)$ the elements are diagonalizable, but not all elements need to have a diagonal representative in their conjugacy class. That is, their diagonal version may lie outside G . A similar consideration holds for G_I .

¹⁷ If $D^R(h)$ has no unit eigenvalue $D^R(h^4)$ could still have it but this can be circumvented by considering another element h^s with suitable real s (analogous to a change in the parameter Y before). What really matters is the uniparametric subgroup $H = \{h^s, s \in \mathbb{R}\}$, or equivalently the Lie algebra generator t of $h = e^t$. Unit eigenvalues of $D^R(h)$ match to zero eigenvalues of t in the representation R .

¹⁸ In the $SU(2)$ example discussed in Sec. IV C, besides the trivial representation, only $j = 1/2$ was involved, so no obstruction arose in that case.

real. Choosing a complex axis¹⁹ would not help though: if J_n has a zero eigenvalue whenever \hat{n} is real (and so $\det(\hat{n} \cdot J) = 0$) by analytical extension, the zero will persist in the complex case too. In what follows we stick to the choice $h \in G_I$.

It follows that for certain groups and representations there is no perfect choice of a single h that would work simultaneously for all components of a general complex probability P . The obvious solution is to try to decompose P as a sum of terms in such a way that each term can be treated effectively by a different suitable element h :

$$P(g) = 1 + \sum_{k=1}^m P_k(g) \quad (h_k \in G_I \text{ and acts effectively on } P_k). \quad (5.2)$$

Eq. (5.1) would then apply for each term $k = 1, \dots, m$ separately without obstruction, and each h_k would introduce a further pair of branches in the support of ρ . The arguments given at the end of Sec. III A indicate the number m of terms should be as small as possible.

In a setting like that of Eq. (4.59), i.e., a matrix group with diagonal h , the obstruction appears for those components with $\Omega = 1$.²⁰ A simpleminded approach would be to use such diagonal h for the $\Omega \neq 1$ terms and a different element h' for the remainder. However such strategy is not practical in general. To see this consider again $SU(2)$ and a representation $R = j$, with integer j (since the half-integer irreps pose no problem). A diagonal $h = h_z$ corresponds to a rotation around the z axis. The components in P can be decomposed in the J_z basis $|j, m\rangle_z$, and $|j, 0\rangle_z$ will be unaffected by h_z . A simple prescription is to identify such components from the condition $\Omega = 1$. All the $\Omega \neq 1$ terms can be treated with h_z (of sufficient magnitude to guarantee positivity of the representation). The terms with $\Omega = 1$ should be treated with a different element h_n , corresponding to a rotation around some axis \hat{n} . As it turns out, one cannot take just any axis. The reason is that we need h_n to act effectively on $|j, 0\rangle_z$: This vector can be decomposed in the basis $|j, m\rangle_n$ and one should take \hat{n} in such a way that $|j, 0\rangle_z$ has no component along $|j, 0\rangle_n$ (since such component would remain unaffected by h_n). Hence, the axis \hat{n} must fulfill the condition

$${}_z\langle j, 0 | j, 0 \rangle_n = 0. \quad (5.3)$$

In practice, this means that the cosine of the angle between the z axis and \hat{n} should be a zero of the j -th Legendre polynomial, $P_j(\hat{e}_z \cdot \hat{n}) = 0$. For all odd j , $\hat{n} = \hat{e}_x$ suffices. Unfortunately for even j the axis must be changed for different j and in general an infinite number of branches could be required.

So a method is needed to implement Eq. (5.2) using a common (and small) set of branches for all representations simultaneously. This can be done as follows.

Let the set of elements $h_k \in G_I$, $k = 1, \dots, m$, where the number m is to be chosen appropriately for the given group.

For any irrep $R \neq 1$, let V^R be the n_R -dimensional vector space where $D^R(g)$ acts ($n_R = \dim R$). Each h_k defines a *singlet* subspace W_k^R of V^R (which may be $\{0\}$); singlet means that within this subspace h_k acts as the identity operator:

$$W_k^R = \{v \in V^R, D^R(h_k)v = v\}. \quad (5.4)$$

On the orthogonal complement $W_k^{R,\perp}$ the element h_k acts effectively [i.e., no non null vector of $W_k^{R,\perp}$ is left invariant by $D^R(h_k)$] and

$$V^R = W_k^R \oplus W_k^{R,\perp}. \quad (5.5)$$

The obstruction is avoided for the irrep R if any vector of V^R can be decomposed as a sum where each term is acted effectively by h_k , i.e.

$$\forall v \in V^R \quad v = \sum_{k=1}^m v_k, \quad v_k \in W_k^{R,\perp}. \quad (5.6)$$

In other words,

$$V^R = W_1^{R,\perp} + \dots + W_m^{R,\perp}. \quad (5.7)$$

(This is the plain sum of subspaces, no mutual null intersection nor orthogonality is assumed.) If Eq. (5.7) holds for a *fixed set of h_k common to all irreps R* , the complex probability representation problem is solved for the group. Note that the P^R appearing in the decomposition of P are matrices rather than vectors of V^R , however, since h acts on the left [e.g. Eq. (5.1)] one can view P^R as a set of n_R column vectors of V^R and apply the method to these vectors, then P^R gets decomposed as a sum of matrices each one acted effectively by one of the h_k , as required in Eq. (5.2). The decomposition $v = \sum_k v_k$ is not unique in general and so some canonical prescription can be adopted to fix the ambiguity.

Now let us show that suitable sets of elements $\{h_k \in G_I, k = 1, \dots, m\}$ do exist for any Lie group G . Let us write $h_k = e^{it_k}$ where t_k are in the Lie algebra of G . A *sufficient* condition to fulfill Eq. (5.7) simultaneously for all irreps R is that the t_k generate the Lie algebra, or equivalently, the elements e^{t_k} generate G .²¹ To see that this is sufficient, let us first note that the condition Eq. (5.7) is equivalent to

$$W_1^R \cap \dots \cap W_m^R = \{0\}. \quad (5.8)$$

This follows from the property $(A + B)^\perp = A^\perp \cap B^\perp$ and the fact that the spaces are finite-dimensional [34]. The equivalence implies that (upon suitable decomposition) the set of elements h_k acts *effectively* on any vector of V^R [Eq. (5.7)] if and only if there are no nontrivial singlet vectors common to all the h_k simultaneously [Eq. (5.8)]. But the latter condition is guaranteed if the e^{t_k} generate G . Indeed, let us assume that

¹⁹ For $g_\pm \in G^c$ Eq. (4.25) generalizes Eq. (4.31).

²⁰ Throughout Ω denotes a generic argument of the function χ , e.g. in Eq. (4.59). Ω is any of the eigenvalues of $D^R(h)$.

²¹ I.e., the minimal algebra containing $\{t_k, k = 1, \dots, m\}$ is the whole algebra, and the minimal subgroup containing all the subgroups $\{e^{st_k}, s \in \mathbb{R}\}$ is G itself.

there were a non trivial singlet $|s\rangle$ common to all the t_k , i.e., $D^R(h_k)|s\rangle = |s\rangle$. Then the stability group of $|s\rangle$ would contain all the e^{t_k} and so it would coincide with G . This would imply that V^R contains a proper invariant subspace (namely the multiples of $|s\rangle$) in contradiction with the assumption that $R \neq 1$ is irreducible.

We have just shown that if the set $\{t_k, k = 1, \dots, m\}$ generates the whole Lie algebra, any P^R can be decomposed as a sum of terms in such a way that at least one of the h_k acts effectively on each term, and this for all the irreps R except the trivial one. Certainly, if one takes as t_k all the elements of a linear basis of the algebra, they generate the whole algebra, so it is never necessary to take m larger than n (n being the dimension of the group G) and in general a smaller m is sufficient.

The condition that the set of elements t_k must generate the whole algebra is sufficient but certainly not necessary in general. Again this is clear in the Abelian case $U(1)^{\times n}$. There, only a whole basis of the algebra would generate the full algebra (and so $m = n$) yet, $m = 1$ is enough as follows from our discussion in Sec. III B 1: A single displacement $h = -i\mathbf{Y}$ with irrational components (so that $\mathbf{k} \cdot \mathbf{Y} \neq 0$) suffices to have an effective action on all Fourier modes simultaneously.

For the general non Abelian case the analysis is more complicated so we stick to our criterion of the set $\{t_k, k = 1, \dots, m\}$, generating the whole Lie algebra. Here we find the remarkable result that for semisimple Lie algebras, $m = 2$ seems to be always sufficient.

For instance, for $SU(2)$ one can take $m = 2$, with $t_1 = i\sigma_z$ and $t_2 = i\sigma_x$.²² V^j has dimension $2j + 1$, the singlet spaces W_1^j and W_2^j have both dimension 1 for integer j or 0 for half-integer j . In both cases $W_1^{j,\perp} + W_2^{j,\perp}$ fills the space V^j . A canonical prescription to decompose $v = v_1 + v_2$, $v \in V^j$, $v_k \in W_k^{j,\perp}$, is to require $v_1 = v_2$ along $W_1^{j,\perp} \cap W_2^{j,\perp}$. This fixes $v_{1,2}$ uniquely. So a total of 4 branches suffice for any complex probability defined on $SU(2)$.

For $SU(3)$ the whole algebra is generated by $i\lambda_2$ and $i\lambda_4$: by taking commutators recursively, eventually a basis of $\mathfrak{su}(3)$ is produced. So four branches suffice also in this case.

Moreover, the following two Lie algebra elements seem to generate the full algebra $\mathfrak{su}(N)$ for any N :

$$\begin{aligned} t_1 &= i \text{diag}(1, 2, \dots, N-1, -N(N-1)/2) \\ (t_2)_{\alpha\beta} &= \begin{cases} i, & \alpha = 1, \beta \neq 1 \text{ or } \beta = 1, \alpha \neq 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (5.9)$$

While we have no rigorous proof of this for all N , the statement holds, at least, for $N \leq 8$. In fact, almost any pair of random elements seem to generate $\mathfrak{su}(N)$, and a smaller sub-

algebra would only be generated by a careful choice of the pair (t_1, t_2) .

The fact that a generic pair of elements $t_{1,2}$ generate the whole algebra is consistent with $\mathfrak{su}(N)$ being simple. As for the direct sum of simple algebras (semisimple algebras), $m = 2$ would hold too. For instance, for $G = G_1 \times G_1$ with $G_1 = SU(2)$. The algebra has basis $i\sigma_j^r$, with $j = 1, 2, 3$, and $r = 1, 2$. It is straightforward to check that the pair of elements $t_x = i\sigma_x^1 + \alpha_x i\sigma_x^2$ and $t_y = i\sigma_y^1 + \alpha_y i\sigma_y^2$ generates $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ for almost any choice of the real coefficients $\alpha_{x,y}$.

If Abelian sectors are added to the semisimple algebra, still $m = 2$ is sufficient to generate the full algebra if the Abelian sector is at most two-dimensional, but not in general. This does not imply though that $m > 2$ is mandatory to fulfill Eq. (5.7), as already shown for the purely Abelian case.

Another remark is that for a higher-dimensional system, with $G = G_1 \times \dots \times G_1$ (n factors), four branches (from $m = 2$) may not be optimal, in the same way that using strictly two-branches (by taking an irrational \mathbf{Y}) is not optimal in the Abelian case. Also in the nonabelian case an uniformity criterion with respect to the n variables is desirable. The same ideas given in Sec. III B 2 apply here, i.e., a bifurcation for each variable and for each of the m terms. So, the number of branches changes from $2m$ to $m2^n$.

B. Case study I

The $SU(2)$ example discussed in Sec. IV C does not contain integer representations, besides the trivial one, and so the problem of a singular kernel does not arise. In order to illustrate the treatment of singular kernels discussed in the previous subsection, let us consider the following “complex” density defined in $SU(2)$,

$$P(g) = \text{tr}(g)^2. \quad (5.10)$$

This probability contains components $j = 0, 1$. It should be noted that actually P is already real and positive, and normalized, but it needs at least four branches in the complexified group if one insists on prescribing a certain decomposition and requires positivity of each component separately.

The density can be written as $P(g) = g^i_i g^j_j$. In order to separate the trivial representation, we can exploit the relation

$$1 = \det(g) = g^1_1 g^2_2 - g^1_2 g^2_1, \quad (5.11)$$

to write

$$P(g) = 1 + g^1_1 g^1_1 + g^2_2 g^2_2 + (g^1_1 g^2_2 + g^1_2 g^2_1), \quad (5.12)$$

corresponding to the decomposition $P = \sum_{jm} P_{jm}$,

$$P(g) = P_{0,0} + P_{1,1} + P_{1,-1} + P_{1,0}. \quad (5.13)$$

The normalization $P_{0,0} = 1$ is to be distributed among the three non trivial components after they are moved into the complexified group manifold.

²² This is not in contradiction with our previous remarks around Eq. (5.3). If $W_{z,n}$ denote the singlet spaces for rotations generated by J_z and J_n respectively, Eq. (5.3) expresses the condition that $W_z \subset W_n^\perp$. This is more restrictive than $V^j = W_z^\perp + W_n^\perp$, V^j being the $2j + 1$ -dimensional space carrying the $SU(2)$ representation j . The condition $W_z \subset W_n^\perp$ does require to change \hat{n} for different j , whereas $V^j = W_z^\perp + W_n^\perp$ does not.

In a first step we can take a diagonal element $h_z \in \text{SU}(2)_I$, corresponding to an imaginary rotation

$$h_z = \begin{pmatrix} \omega_z & 0 \\ 0 & \omega_z^{-1} \end{pmatrix}, \quad \omega_z > 1, \quad (5.14)$$

which would produce [using Eq. (4.56)]

$$\begin{aligned} \hat{Q}_{1,1} + \hat{Q}_{1,-1} &= g^1_1 g^1_1 \chi(\omega_z^2) + g^2_2 g^2_2 \chi(\omega_z^{-2}), \\ \hat{Q}_{1,0} &= (g^1_1 g^2_2 + g^1_2 g^2_1) \chi(1). \end{aligned} \quad (5.15)$$

The terms $|1, \pm 1\rangle$ can be treated with h_z but $|1, 0\rangle$ requires a different transformations since it is invariant under rotations around the z axis and $\chi(1)$ diverges.

For $P_{1,0}$ one can apply a rotation around the x axis relying on $\langle 1, 0 | \hat{R}(\hat{e}_x, \pi/2) | 1, 0 \rangle = 0$,

$$h_n = U h_x U^{-1}, \quad h_x = \text{diag}(\omega_x, \omega_x^{-1}), \quad U = e^{-i\pi\sigma_y/4}. \quad (5.16)$$

An alternative to computing the rank four tensor $\chi(h_n)^{i_1 i_2}_{j_1 j_2}$ is to rotate the elements, as explained in Sec. IV E: the effect of h_n on g corresponds to the action of h_x on $g' = U^{-1} g U$. Since h_x is diagonal Eq. (4.56) applies. The explicit result in terms of g' becomes

$$\begin{aligned} \hat{Q}_{1,0} &= \frac{1}{2} ((g'^1_1)^2 \chi(\omega_x^2) - (g'^1_2)^2 \chi(\omega_x^{-2}) \\ &\quad - (g'^2_1)^2 \chi(\omega_x^2) + (g'^2_2)^2 \chi(\omega_x^{-2})). \end{aligned} \quad (5.17)$$

As advertised no divergence of the type $\chi(1)$ arises.

After this decomposition the expectation values can be expressed through real weights on the complexified group with four sheets

$$\langle A \rangle = \int_{\text{SU}(2)} dg \sum_{\sigma=\pm 1} (Q_{z,\sigma}(g) A(g h_z^\sigma) + Q_{x,\sigma}(g) A(g h_n^\sigma)). \quad (5.18)$$

Following Eq. (4.53), here $Q_{z,+}$ is twice the real part of $\hat{Q}_{1,1} + \hat{Q}_{1,-1}$ plus some constant term $N_{z,+}$ from $P_{0,0}$, $Q_{z,-}$ is likewise with h_z^{-1} , and $Q_{x,\pm}$ likewise for $\hat{Q}_{1,0}$ with h_n . The positive constant terms $N_{z,\pm}$, $N_{x,\pm}$ add up to one.

In our case, the two functions $Q_{z,\pm}$ turn out to be equal, after choosing equal normalizations $N_{z,+} = N_{z,-}$, and similarly for $Q_{x,\pm}$. An explicit calculation gives

$$\begin{aligned} Q_z(g) &= N_z + 2 \frac{\cos^2(\psi/2) - \cos^2(\theta) \sin^2(\psi/2)}{\omega_z^2 + \omega_z^{-2}}, \\ Q_x(g) &= N_x + \frac{\cos^2(\theta) + \cos(\psi) \sin^2(\theta)}{\omega_x^2 + \omega_x^{-2}}, \end{aligned} \quad (5.19)$$

with

$$N_z, N_x \geq 0, \quad 2N_z + 2N_x = 1. \quad (5.20)$$

In the formulas $g = e^{-i\psi \hat{n} \cdot \sigma/2}$ and $\hat{n} = (\theta, \phi)$ in spherical coordinates. ϕ does not appear in our case, related with the invariance of $P(g)$ with respect to similarity transformations of g .

Upon minimization with respect to (θ, ψ) , the conditions ensuring positive functions $Q_z(g)$ and $Q_x(g)$ are

$$\begin{aligned} 0 &\leq \min Q_z = N_z - 2(\omega_z^2 + \omega_z^{-2})^{-1}, \\ 0 &\leq \min Q_x = N_x - (\omega_x^2 + \omega_x^{-2})^{-1}. \end{aligned} \quad (5.21)$$

These inequalities can be fulfilled by taking $\omega_{z,x}$ sufficiently large. The optimal case (smaller $\omega_{z,x}$) corresponds to $\min Q_z = \min Q_x = 0$, i.e.,

$$\begin{aligned} 2(\omega_z^2 + \omega_z^{-2})^{-1} &= N_z, \quad (\omega_x^2 + \omega_x^{-2})^{-1} = \frac{1}{2} - N_z, \\ 0 &\leq N_z \leq \frac{1}{2}. \end{aligned} \quad (5.22)$$

For instance, for $N_z = 1/3$ one obtains $\omega_z = \omega_x = 1 + \sqrt{2}$, while for $N_z = 1/4$, $\omega_z = 2.81$ and $\omega_x = 1.93$.

Formally it would seem that one could remove, say the two sheets $Q_{z,\pm}(g)$ by taking $\omega_z \rightarrow \infty$ and $\omega_x \rightarrow 1$ [or $Q_{x,\pm}(g)$ with $\omega_z \rightarrow 1.93$ and $\omega_x \rightarrow \infty$] however, this is incorrect. For large ω_z , $Q_{z,\pm}(g)$ is reduced but the information must be carried by the observable, $A(g h_z^{\pm 1})$. The observables tend to grow rapidly far from the real manifold producing an infinite variance in the limit.

It is noteworthy that the functions $Q_{z,x}(g)$ in Eq. (5.19) do not diverge as $\omega_{z,x} \rightarrow 1$. This is a consequence of the fact that our $P(g)$ is real. In that limit the four distributions have their support on the real manifold and their sum reproduces the original density:

$$h_z = h_x = 1 : \quad 2Q_z(g) + 2Q_x(g) = 2(1 + \cos \psi) = P(g). \quad (5.23)$$

Even if in the limit $\omega_{z,x} = 1$ the sum of the four contributions yield the original positive density, Q_z and Q_x would not be separately positive. It is the requirement $Q_z(g) \geq 0$ and $Q_x(g) \geq 0$ that introduces the non trivial lower bounds on ω_z and ω_x .

C. Case study II

Next we consider a complex probability defined on $G = \text{SU}(N) \times \text{SU}(N)$, representing a simplified lattice with two degrees of freedom, namely,

$$\begin{aligned} P(g_1, g_2) &= \mathcal{N}^{-1} (1 + \beta \text{tr}(g_1^{-1} g_2)) (1 + \beta \text{tr}(g_2^{-1} g_1)) \\ &\quad \times \text{tr}(g_1) \text{tr}(g_2^{-1}). \end{aligned} \quad (5.24)$$

The terms with β mimic a gauge action. Those factors are invariant under $g_i \rightarrow \omega^{-1} g_i \omega'$, $i = 1, 2$, $\omega, \omega' \in \text{SU}(N)$. The factors $\text{tr}(g_1) \text{tr}(g_2^{-1})$ mimic Polyakov loops, partially breaking the invariance from $\text{SU}(N) \times \text{SU}(N)$ to $\text{SU}(N)$ ($\omega = \omega'$), but preserving global center invariance, $g_i \rightarrow z g_i$, $z \in \text{U}(1)$, $z^n = 1$.

For $N > 2$ the normalization of $P(g)$ comes solely from $\text{tr}(g_1^{-1} g_2) \text{tr}(g_1) \text{tr}(g_2^{-1})$, however when $N = 2$ the term

$\text{tr}(g_2^{-1}g_1)\text{tr}(g_1)\text{tr}(g_2^{-1})$ gives an identical contribution, due to $\text{tr}(g^{-1}) = \text{tr}(g)$. Thus $P(g)$ is normalized with²³

$$\mathcal{N} = \begin{cases} \beta & (N=2) \\ \beta/N & (N>2) \end{cases}. \quad (5.25)$$

One can decompose $P(g)$ in monomials, as in Eq. (4.60), and apply a diagonal element of G_I, h_z , with parameters $\tilde{\omega}_{z,i}^r > 0, r=1,2, i=1,\dots,N$. The complex representation \hat{Q} is then obtained as in Eq. (4.61). Each term in \hat{Q} picks up a factor $\chi(\Omega_z)$ and the problem of singular kernels corresponds to the components for which $\Omega_z = 1$. Such components should be treated with a different element h_n of G_I .

We can see that the terms which are singular under h_z , i.e., contain the trivial representation (in a reduction with respect to the subgroup generated by h_z) are contained in $\text{tr}(g_1^{-1}g_2)\text{tr}(g_1)\text{tr}(g_2^{-1})$.

$$\begin{aligned} \text{tr}(g_1^{-1}g_2)\text{tr}(g_1)\text{tr}(g_2^{-1}) &= g^{-1i}{}_j g^j{}_i g^k{}_k g^{2-1\ell}{}_j, \\ \Omega_z &= \tilde{\omega}_{z,i}^{-1} \tilde{\omega}_{z,i}^2 \tilde{\omega}_{z,k}^1 \tilde{\omega}_{z,\ell}^{-1}. \end{aligned} \quad (5.26)$$

Generically $\Omega_z = 1$ when $i = k = \ell$, a total of N^2 terms:

$$g^1{}_i g^{-1i}{}_j g^2{}_j g^{-2i}{}_i, \quad i, j = 1, \dots, N, \quad \Omega_z = 1. \quad (5.27)$$

In order to choose h_n , this can be analyzed as follows. Each factor $\tilde{g}^r \otimes \tilde{g}^{-r}$, $r=1,2$, can be reduced as trivial plus adjoint representation and contains N singlets under a diagonal \tilde{h} (one from the trivial representation and $N-1$ from the adjoint). This N -dimensional space is spanned by the $N \times N$ diagonal matrices (the traceless matrices being in the adjoint sector). Therefore, out of the N^2 singular terms, one comes from the trivial representation of $\text{SU}(N) \times \text{SU}(N)$ and the remaining $N^2 - 1$ come from the adjoint representation in one or both factors. So h_n can be chosen in the form $\tilde{h}_n^1 \tilde{h}_n^2$ with the condition that \tilde{h}_n must act *effectively* on the components of $\tilde{g}^r \otimes \tilde{g}^{-r}$ which are invariant under h_z . If h_n is written as $U h_x U^{-1}$, with diagonal h_x , U must be chosen so that any traceless diagonal matrix, upon rotation by U , has not overlap with any other traceless diagonal matrix (similar to the condition in Eq. (5.3)):

$$\begin{aligned} 0 &= \text{tr}(A_z U A_x U^{-1}), \quad U \in \text{SU}(N) \\ &\text{for all } A_{z,x} \text{ traceless and diagonal.} \end{aligned} \quad (5.28)$$

An easy calculation shows that this implies

$$|U^j{}_\ell|^2 = \frac{1}{N} \quad j, \ell = 1, \dots, N, \quad (5.29)$$

and an explicit solution is

$$U^j{}_\ell = \frac{1}{\sqrt{N}} e^{i2\pi(j-1)(\ell-1)/N} \quad j, \ell = 1, \dots, N. \quad (5.30)$$

In particular for $N=2$, $U = e^{-i\pi\sigma_y/4}$ [consistently with Eq. (5.16)].

One can now verify that the previously singular terms of Eq. (5.27) are not singular under h_n [upon removing the trivial representation of $\text{SU}(N) \times \text{SU}(N)$]. To do that we use

$$g^i{}_j \mapsto (gh_n)^i{}_j = (gU h_x U^{-1})^i{}_j = \sum_\ell (gU)^i{}_\ell \omega_{x,\ell} (U^{-1})^\ell{}_j. \quad (5.31)$$

It is sufficient to consider just one of the factors in (5.27):

$$\begin{aligned} g^1{}_i g^{-1i}{}_j &\mapsto \sum_\ell (gU)^i{}_\ell \tilde{\omega}_{x,\ell}^1 (U^{-1})^\ell{}_i \sum_m U^i{}_m \tilde{\omega}_{x,m}^{-1} (U^{-1})^m{}_j, \\ \Omega_x &= \tilde{\omega}_{x,\ell}^1 \tilde{\omega}_{x,m}^{-1}. \end{aligned} \quad (5.32)$$

The possible singular contributions, $\Omega_x = 1$, would come from $\ell = m$. For these terms one obtains

$$\sum_\ell (gU)^i{}_\ell (U^{-1})^\ell{}_i U^i{}_\ell (U^{-1})^\ell{}_j = \frac{1}{N} \delta^i{}_j, \quad (5.33)$$

using Eq. (5.29).²⁴ An identical result is obtained for the second factor $\tilde{g}^2{}_j \tilde{g}^{-2i}{}_i$. So the terms that remain invariant under h_n are

$$g^1{}_i g^{-1i}{}_j g^2{}_j g^{-2i}{}_i \mapsto \frac{1}{N} \delta^i{}_j \frac{1}{N} \delta^j{}_i = \frac{1}{N}. \quad (5.34)$$

This is independent of g and corresponds to the trivial representation of the full group, which always has to be extracted from $P(g)$. The trivial representation saturates the normalization, and indeed, the final result $1/N$ combined with the factor $\mathcal{N}^{-1}\beta$ (or $2\mathcal{N}^{-1}\beta$ for $N=2$) checks that $P(g)$ is normalized.

After extraction of the constant mode, $P(g)$ can be written as a sum of two terms, namely, the monomials to be rotated with h_z and those to be rotated with h_n ,

$$P(g) = 1 + P_z(g) + P_x(g). \quad (5.35)$$

It should be noted that P_z (the same goes for P_x) is non singular for generic values of $\tilde{\omega}_{z,j}$, but new divergences can appear for especial correlated values. For instance a term with $\chi(\tilde{\omega}_{z,i}^1 \tilde{\omega}_{z,j}^{-1})$ prevents taking these two ω 's to be equal.

Let us consider the case $N=2$ in more detail:

$$\begin{aligned} P(g) &= \frac{1}{\beta} (1 + \beta \text{tr}(g_1^{-1}g_2))^2 \text{tr}(g_1)\text{tr}(g_2), \\ g &= (g_1, g_2) \in \text{SU}(2) \times \text{SU}(2). \end{aligned} \quad (5.36)$$

In addition, for simplicity, we will assume $\beta > 0$.

²³ Using standard $\text{SU}(N)$ group integration rules [35].

²⁴ Alternatively, one can derive the condition in Eq. (5.29) by requiring the fulfillment of (5.33).

The complex representations associated to the two sectors P_z and P_x are easily obtained using $\hat{h}_z = \text{diag}(\hat{\omega}_z, \hat{\omega}_z^{-1})$, and similarly for \hat{h}_x . This gives [expanding $P_{z,x}$ in monomials and applying Eq. (4.56)]

$$\begin{aligned}\hat{Q}_z &= 2g^1_1 g^2_1 \left(g^1_1 g^2_2 - g^2_1 g^1_2 \right) \chi(\omega_z^2) + \dots \quad (16 \text{ terms}) \\ \hat{Q}_x &= \frac{1}{2} \left(g^1_1 + g^1_2 \right) \left(g^2_1 + g^2_2 \right) \chi(\omega_x^2) + \dots \quad (8 \text{ terms})\end{aligned}\quad (5.37)$$

The 16 terms are classified by 16 combinations of the exponents (k, m) in $\chi(\omega_z^k \omega_z^m)$, and similarly for the 8 terms in the x sector.

Taking real parts, and changing $\omega \rightarrow \omega^{-1}$, for the various ω , produces the four distributions corresponding to four sheets on the complexified group, two sheets for each sector z and x . After this step the dependence on the Ω_z 's is through the symmetric combination $\chi(\Omega_z) + \chi(\Omega_z^{-1})$, and similarly in the x sector. This feature is an idiosyncrasy of this complex probability and group.

However, as discussed in Sec. III B 2, instead of two sheets, it is preferable to use 2^n sheets for n variables, $n = 2$ in our case. This allows to take the same $\hat{\omega}_z$ for $r = 1$ and $r = 2$ (and similarly for $\hat{\omega}_x$), and also to reduce the numerical value of the $\Omega_{z,x}$ required to have positive distributions.

The method is explained in Sec. III B 2: Initially there are two sheets in the z sector (everything is similar in the x sector), produced by the transformations $(\hat{\omega}_z, \hat{\omega}_z)$ and $(\hat{\omega}_z^{-1}, \hat{\omega}_z^{-1})$. Then a term with $\Omega_z = \hat{\omega}_z^k \hat{\omega}_z^m$ is unchanged if $km > 0$. If $km < 0$, it is changed to $\hat{\omega}_z^k \hat{\omega}_z^{-m}$ and moved to the sheet $(\hat{\omega}_z, \hat{\omega}_z^{-1})$. When $km = 0$, half of the term stays and the other half is moved to the opposite sheet.²⁵

Following this procedure eight branches, with functions $Q_{z\pm,\pm}(g)$ and $Q_{x\pm,\pm}(g)$, are obtained. Taking the symmetric choice

$$\hat{\omega}_z = \hat{\omega}_z \equiv \omega_z, \quad \hat{\omega}_x = \hat{\omega}_x \equiv \omega_x, \quad (5.38)$$

Q_{z++} contains terms $\Omega_z = \omega_z^m$ with $m = 2$, while Q_{z+-} has $m = 2, 4, 6$. In the x sector, Q_{x++} and Q_{x+-} both contain terms $\Omega_x = \omega_x^m$ with $m = 2, 4$.

In order to apply the method, the unit normalization of P must be distributed among the eight branches to produce positive distributions. To achieve this $\omega_{z,x}$ have to be taken sufficiently large so that all minima of $Q_{z,\pm,\pm}$ and $Q_{x,\pm,\pm}$, and their sums, are above -1 .²⁶ The minima of these functions (over the manifold $\text{SU}(2) \times \text{SU}(2)$) will depend on the choice

of $\omega_{z,x}$ and β and presumably they cannot be found in a closed analytical form. Our approach has been to split the functions into a sum of terms classified by their dependence on Ω and its power of β and (numerically) find an independent minimum for each such term. This provides a lower bound to the true minimum, since there can be cancellations between terms which are neglected in our approach. A lower bound is sufficient for our purposes. The lower bounds to the minima so obtained are

$$\begin{aligned}\min Q_{z,++} &\geq -(4\beta + 4 + \frac{2}{\beta})\chi_s(\omega_z^2) \\ \min Q_{z,+ -} &\geq -(6\beta + 4 + \frac{2}{\beta})\chi_s(\omega_z^2) \\ &\quad - (4\beta + 4)\chi_s(\omega_z^4) - 2\beta\chi_s(\omega_z^6) \\ \min Q_{x,++} &\geq -\chi_s(\omega_x^2) - \chi_s(\omega_x^4) \\ \min Q_{x,+ -} &\geq -\chi_s(\omega_x^2) - \chi_s(\omega_x^4)\end{aligned}\quad (5.39)$$

where $\beta > 0$ and

$$\chi_s(\Omega) \equiv \chi(\Omega) + \chi(\Omega^{-1}) = \frac{1}{\Omega + \Omega^{-1}}. \quad (5.40)$$

It is noteworthy that the coefficients found numerically turn out to be simple numbers. Remarkably, choosing concrete values of β (to combine various terms and so increase the minimum) has not resulted in any improvement. So the method used seems to be numerically accurate, producing good estimates for the minima.

Since all expressions in Eq. (5.40) are negative, it is sufficient to constrain their sum. The optimal values of the pair (ω_z, ω_x) are thus constrained by the condition

$$\begin{aligned}\frac{1}{2} &= (10\beta + 8 + \frac{4}{\beta})\chi_s(\omega_z^2) + (4\beta + 4)\chi_s(\omega_z^4) \\ &\quad + 2\beta\chi_s(\omega_z^6) + 2\chi_s(\omega_x^2) + 2\chi_s(\omega_x^4).\end{aligned}\quad (5.41)$$

Saturation of the equality by the terms with ω_z (by letting ω_x to be as high as needed) yields the bounds $\omega_z \geq e^{1.90}$ for $\beta = 1$ and $\omega_z \geq e^{2.05}$ for $\beta = 2$. Likewise $\omega_x \geq e^{0.72}$ for any value of β .

The choice $\omega_z = \omega_x$, for $\beta = 1$ and $\beta = 2$ gives $\omega_z = 6.95 = e^{1.94}$ and $\omega_z = 8.02 = e^{2.08}$, respectively. In this scenario most of the normalization (92%) goes to the z -sheets, with $N_{z,++} = 0.207$, $N_{z,+ -} = 0.251$, and $N_{x,++} = N_{x,+ -} = 0.021$. Using these parameters, we have analyzed a sample operator, $\mathcal{O} = \text{tr}(g_1^{-1} g_2)$, with exact expectation value is $\langle \mathcal{O} \rangle = \beta + 1/(2\beta)$. In our representation, the expectation value comes only from the sheets $Q_{z,+ -}$ and $Q_{z,- +}$, the other sheets giving a vanishing contribution. All the sheets contribute to the variance, which can be computed analytically, but $Q_{z,+ -}$ and $Q_{z,- +}$ are also dominant for the variance, through a large β -independent term, namely, $N_{z,+ -} \omega_z^4$. For $\beta = 1$ the total variance is 634.²⁷

²⁵ The coordinate that is reflected is that with a zeroth power in Ω_z .

²⁶ Here the functions $Q_{z,x}$ do not contain the constant modes. The conditions to be above -1 are similar to those in Eq. (3.10). They guarantee that a global unit normalization can be added to the various branches in the form constant modes to make these functions positive.

²⁷ The number depends also on the precise definition of the variance. The

This is to be compared with the variance obtained using simple reweighting with $|P(g)|$ (and assuming that its normalization is known). This variance can be obtained analytically, obtaining

$$\text{Var}_{\text{RW}} = 0.878 + 0.374\beta^2 + \frac{0.331}{\beta^2}. \quad (5.42)$$

This gives number of the order of unity for β 's of the same order. Therefore in this case reweighting has a much better performance than the representation, however such good performance should deteriorate exponentially with the number of variables.

VI. SUMMARY AND CONCLUSIONS

In this work we have analyzed the problem of constructing representations of complex weights within the two-branch approach, which is optimal from the point of view of localization. In the Abelian many-dimensional case a solution is found to the problem of treating all variables on an equal footing, and simplifying the choice of parameters. The method proposed is to share the weight over 2^n sheets, for n variables. This allows to use copies of the real manifold which are closer (to the real manifold), and so with smaller variance.

The main novelty is the study of representations of complex weight defined on compact group manifolds, within a two-branch approach. In this scheme two copies of the (real) group are obtained upon translation by an imaginary element and its inverse. Each copy carries a positive distribution whose

analytical continuation, when added, reproduce the original complex weight. The construction is illustrated in detail for a complex weight defined on $SU(2)$. When the imaginary element does not act effectively on some of the components of the complex weight, so that they are not moved to the complexified group manifold, an obstruction is met in the form of a singular kernel. We have shown how the obstruction can be removed, namely, by decomposing the complex weight into components, each of which can be acted effectively by some imaginary element. We have shown that such a decomposition always exists.

Explicit examples have been worked out for $SU(2)$ with integer spin representations (hence, subject to obstruction), and for $SU(2) \times SU(2)$, also presenting singular kernels.

An interesting lesson from the representation approach to the sign problem is that even realistic theories like lattice QCD with a chemical potential must admit such representations (however complicated and nonlocal they might be). This opens the possibility of trying to directly model a local and positive weight on the complexified manifold, incorporating the chemical potential, hopefully in the same universality class as the original QCD problem.

Acknowledgments

I thank E. Seiler and J. Wosiek for discussions. This work has been partially supported by the Spanish Ministerio de Economía y Competitividad (grants Nos. FIS2014-59386-P and FIS2017-85053-C2-1-P) and by the Junta de Andalucía (grant No. FQM-225).

-
- [1] N. Madras, *Lectures on Monte Carlo Methods*, The Fields Institute for Research in Mathematical Sciences, American Mathematical Society, 2002.
 - [2] M. Troyer and U. J. Wiese, *Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations*, Phys. Rev. Lett. **94** (2005) 170201 doi:10.1103/PhysRevLett.94.170201 [cond-mat/0408370].
 - [3] K. Osterwalder and E. Seiler, *Gauge Field Theories on the Lattice*, Annals Phys. **110** (1978) 440. doi:10.1016/0003-4916(78)90039-8
 - [4] P. Hasenfratz and F. Karsch, *Chemical Potential on the Lattice*, Phys. Lett. **125B** (1983) 308. doi:10.1016/0370-2693(83)91290-X
 - [5] O. Philipsen, *Lattice calculations at non-zero chemical potential: The QCD phase diagram*, PoS CONFINEMENT **8** (2008) 011.
 - [6] S. Chandrasekharan and U. J. Wiese, *Meron cluster solution of a fermion sign problem*, Phys. Rev. Lett. **83** (1999) 3116 doi:10.1103/PhysRevLett.83.3116 [cond-mat/9902128].
 - [7] G. Parisi, *On Complex Probabilities*, Phys. Lett. B **131**, 393 (1983).
 - [8] J. R. Klauder, *Stochastic Quantization*, Acta Phys. Austriaca Suppl. **25** (1983) 251.
 - [9] F. Karsch and H. W. Wyld, *Complex Langevin Simulation of the $SU(3)$ Spin Model With Nonzero Chemical Potential*, Phys. Rev. Lett. **55** (1985) 2242. doi:10.1103/PhysRevLett.55.2242
 - [10] G. Aarts, F. A. James, J. M. Pawłowski, E. Seiler, D. Sexty and I. O. Stamatescu, *Stability of complex Langevin dynamics in effective models*, JHEP **1303** (2013) 073 doi:10.1007/JHEP03(2013)073 [arXiv:1212.5231 [hep-lat]].
 - [11] D. Sexty, *Simulating full QCD at nonzero density using the complex Langevin equation*, Phys. Lett. B **729** (2014) 108 doi:10.1016/j.physletb.2014.01.019 [arXiv:1307.7748 [hep-lat]].
 - [12] J. Ambjorn, M. Flensburg and C. Peterson, *The Complex Langevin Equation and Monte Carlo Simulations of Actions With Static Charges*, Nucl. Phys. B **275** (1986) 375. doi:10.1016/0550-3213(86)90605-X
 - [13] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, *Complex Langevin: Etiology and Diagnostics of its Main Problem*, Eur. Phys. J. C **71** (2011) 1756 doi:10.1140/epjc/s10052-011-1756-5 [arXiv:1101.3270 [hep-lat]].
 - [14] L. L. Salcedo, *Spurious solutions of the complex Langevin equation*, Phys. Lett. B **305** (1993) 125. doi:10.1016/0370-2693(93)91116-5
 - [15] L. L. Salcedo, *Does the complex Langevin method give unbiased results?*, Phys. Rev. D **94** (2016) no.11, 114505 doi:10.1103/PhysRevD.94.114505 [arXiv:1611.06390 [hep-lat]].
 - [16] E. Seiler, *Status of Complex Langevin*, EPJ Web Conf. **175** (2018) 01019 doi:10.1051/epjconf/201817501019 [arXiv:1708.08254 [hep-lat]].

- [17] M. Cristoforetti F. Di Renzo and L. Scorzato [AuroraScience Collaboration], *New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble*, Phys. Rev. D **86** (2012) 074506 doi:10.1103/PhysRevD.86.074506 [arXiv:1205.3996 [hep-lat]].
- [18] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway and N. C. Warrington, *Sign problem and Monte Carlo calculations beyond Lefschetz thimbles*, JHEP **1605** (2016) 053 doi:10.1007/JHEP05(2016)053 [arXiv:1512.08764 [hep-lat]].
- [19] P. F. Bedaque, *A complex path around the sign problem*, EPJ Web Conf. **175** (2018) 01020 doi:10.1051/epjconf/201817501020 [arXiv:1711.05868 [hep-lat]].
- [20] L. L. Salcedo, *Representation of complex probabilities*, J. Math. Phys. **38**, 1710 (1997).
- [21] D. Weingarten, *Complex probabilities on R^{**N} as real probabilities on C^{**N} and an application to path integrals*, Phys. Rev. Lett. **89** (2002) 240201 doi:10.1103/PhysRevLett.89.240201 [quant-ph/0210195].
- [22] L. L. Salcedo, *Existence of positive representations for complex weights*, J. Phys. A **40**, 9399 (2007).
- [23] J. Wosiek, *Beyond complex Langevin equations I: two simple examples*, arXiv:1511.09083 [hep-lat].
- [24] J. Wosiek, *Beyond complex Langevin equations: from simple examples to positive representation of Feynman path integrals directly in the Minkowski time*, JHEP **1604** (2016) 146 doi:10.1007/JHEP04(2016)146 [arXiv:1511.09114 [hep-th]].
- [25] L. L. Salcedo, *Gibbs sampling of complex valued distributions*, Phys. Rev. D **94** (2016) no.7, 074503.
- [26] J. Wosiek, *Beyond complex Langevin equations*, PoS LATTICE **2016** (2017) 314 [arXiv:1610.06874 [hep-lat]].
- [27] E. Seiler and J. Wosiek, *Positive Representations of a Class of Complex Measures*, J. Phys. A **50** (2017) no.49, 495403 doi:10.1088/1751-8121/aa9310 [arXiv:1702.06012 [hep-lat]].
- [28] E. Seiler and J. Wosiek, *Beyond Complex Langevin Equations: positive representation of a class of complex measures*, EPJ Web Conf. **175** (2018) 11004 doi:10.1051/epjconf/201817511004 [arXiv:1710.03533 [hep-lat]].
- [29] B. Ruba and A. Wyrzykowski, *Explicit positive representation for weights on R^d* , EPJ Web Conf. **175** (2018) 11022 doi:10.1051/epjconf/201817511022 [arXiv:1710.04318 [hep-lat]].
- [30] A. Wyrzykowski and B. Ruba, *Satisfying positivity requirement in the Beyond Complex Langevin approach*, doi:10.1051/epjconf/201817511026 arXiv:1710.04828 [hep-lat].
- [31] L. L. Salcedo, *Representation of complex probabilities and complex Gibbs sampling*, EPJ Web Conf. **175** (2018) 07037 doi:10.1051/epjconf/201817507037 [arXiv:1710.03195 [hep-lat]].
- [32] K. Okano, L. Schulke and B. Zheng, *Complex Langevin simulation*, Prog. Theor. Phys. Suppl. **111** (1993) 313. doi:10.1143/PTPS.111.313
- [33] G. Aarts, E. Seiler and I. O. Stamatescu, *The Complex Langevin method: When can it be trusted?*, Phys. Rev. D **81** (2010) 054508 doi:10.1103/PhysRevD.81.054508 [arXiv:0912.3360 [hep-lat]].
- [34] P. R. Halmos, *Finite-dimensional vector spaces*, (Springer-Verlag, Berlin, 1974).
- [35] M. Creutz, *Quarks, gluons and lattices*, (Cambridge University Press, Cambridge, 1983).