

Probability density function for neutrino masses and mixingsJean-François Fortin,^{*} Nicolas Giasson,[†] and Luc Marleau[‡]*Département de Physique, de Génie Physique et d'Optique, Université Laval,
Québec City, Québec G1V 0A6, Canada*

(Received 29 September 2016; published 1 December 2016)

The anarchy principle leading to the seesaw ensemble is studied analytically with the usual tools of random matrix theory. The probability density function for the seesaw ensemble of $N \times N$ matrices is obtained in terms of a multidimensional integral. This integral involves all light neutrino masses, leading to a complicated probability density function. It is shown that the probability density function for the neutrino mixing angles and phases is the appropriate Haar measure. The decoupling of the light neutrino masses and neutrino mixings implies no correlation between the neutrino mass eigenstates and the neutrino mixing matrix and leads to a loss of predictive power when comparing with observations. This decoupling is in agreement with some of the claims found in the literature.

DOI: [10.1103/PhysRevD.94.115004](https://doi.org/10.1103/PhysRevD.94.115004)**I. INTRODUCTION**

The Standard Model (SM) is the pinnacle of our understanding of particle physics. It however comes with a plethora of parameters, the masses and the flavor mixings, that are seemingly not fixed by any known fundamental principle. On one hand, the spectrum of quarks spans 5 orders of magnitude and the quark mixing matrix exhibits some form of flavor hierarchy. On the other hand, the spectrum of charged leptons spans 4 orders of magnitude while the neutrinos are strictly massless, hence the lepton mixing matrix is trivial. Observations of neutrino oscillations [1] however cannot be accommodated in the SM and this suggests to extend the SM to allow for neutrino masses and neutrino mixings. Experimental data show that the neutrino sector prefers three massive light neutrinos with different masses and a neutrino mixing matrix exhibiting near-maximal mixing.

To make sense of the neutrino sector, it was argued in [2,3] that the light neutrino mass matrix could be generated randomly from a more fundamental Dirac neutrino mass matrix and a more fundamental Majorana neutrino mass matrix with random elements distributed according to a Gaussian ensemble, a principle dubbed the anarchy principle. These more fundamental neutrino mass matrices would come from the extended SM where the seesaw mechanism occurs. In [3], it was argued that the probability density function (PDF) for the mixing angles and phases is the appropriate Haar measure of the symmetry group, implying near-maximal mixings. Then, the anarchy principle was analyzed mostly numerically in a series of papers [4–6], reaching interesting conclusions, for example about the preferred normal hierarchy of the neutrino masses.

Although several numerical results have been obtained, few analytical results on the seesaw ensemble, which is derived from the anarchy principle, exist. That it is the case even though random matrix theory is a well-studied subject in mathematics [7–9] is surprising. It is therefore clear that a thorough analytical investigation of the seesaw ensemble is possible.

This paper undertakes this task by investigating analytically the seesaw ensemble derived from the anarchy principle with the help of the usual tools of random matrix theory. The seesaw ensemble PDF is obtained from $N \times N$ fundamental Dirac and Majorana neutrino mass matrices with real or complex elements. The result is given in terms of a specific multidimensional integral. The joint PDF for the singular (eigen) values in the complex (real) case is then derived and it is shown that the group variables decouple straightforwardly as in the usual Gaussian ensembles. Simple properties of the seesaw ensemble PDF are also presented and their implications for the physical case of $N = 3$ are briefly discussed. The full investigation of the seesaw ensemble for the physical case of interest, SM neutrino physics with $N = 3$, will be discussed elsewhere [10].

This paper is organized as follows: Sec. II quickly reviews the type I seesaw mechanism. In Sec. III the PDF for the seesaw ensemble is obtained from the PDFs related to the Dirac and Majorana neutrino mass matrices. Section IV gives the general properties of the seesaw ensemble from the PDF obtained previously. A discussion, including a comparison with the existing literature, and a conclusion are presented in Sec. V. Finally, Appendixes A and B derive explicitly some useful results.

II. TYPE I SEESAW MECHANISM

This section reviews the type I seesaw mechanism [11].

^{*}jean-francois.fortin@phy.ulaval.ca

[†]nicolas.giasson.1@ulaval.ca

[‡]luc.marleau@phy.ulaval.ca

A. Type I seesaw mechanism

In the SM, neutrinos, which are part of the doublets L^i where i denotes one of the three families, are massless. To account for the oscillations observed experimentally, the type I seesaw mechanism postulates the existence of three singlet fermions N^i (left-handed Weyl spinors) that play the role of the “missing” right-handed neutrinos $\sigma^2 N^{i*}$. In terms of the SM Lagrangian, the type I theory is described by

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \bar{N}^i i \bar{\sigma}^\mu \partial_\mu N^i - \left(\frac{1}{2} M_{R,ij} N^i N^j + y_{D,ij} L^i H N^j + \text{H.c.} \right), \quad (2.1)$$

where H is the Higgs doublet and the allowed terms correspond to right-handed neutrino masses M_R^{ij} and Yukawa interactions y_D^{ij} . From (2.1) the (νN) neutrino mass matrix is given by

$$M_{\nu N} = \begin{pmatrix} 0 & M_D \\ M_D^T & M_R \end{pmatrix}, \quad (2.2)$$

where $M_D = y_D v / \sqrt{2}$ is the Dirac neutrino mass matrix and v is the Higgs vacuum expectation value, i.e. $\langle H \rangle = \frac{1}{\sqrt{2}} (0 \ v)^T$.

The right-handed neutrino mass scale could naturally be much larger than the Dirac neutrino mass scale since the former is an (*a priori* arbitrary) energy scale while the latter originates from a Yukawa coupling, hence (2.2) implies that the light neutrino mass matrix is

$$M_\nu = -M_D M_R^{-1} M_D^T. \quad (2.3)$$

The (symmetric) neutrino mass matrix (2.3) is the starting point to compute the seesaw ensemble PDF from the PDFs for the Dirac and Majorana neutrino mass matrices M_D and M_R , as described in the next section.

III. DEFINITION OF THE SEESAW ENSEMBLE

In this section we obtain the PDF for the type I seesaw ensemble from the PDFs associated to the Dirac and Majorana neutrino mass matrices. To be as general as possible, the fundamental neutrino mass matrices are taken to be $N \times N$ instead of 3×3 .

A. Dirac and Majorana ensembles

Following the principle of anarchy [3,6], where elements of the fundamental matrices M_D and M_R are random, the PDFs for the Dirac and Majorana neutrino mass matrices are Gaussian and defined as

$$P_D(M_D) dM_D = \left(\frac{1}{2\pi\Lambda_D^2} \right)^{\beta N^2/2} \exp \left[-\frac{\text{tr}(M_D^\dagger M_D)}{2\Lambda_D^2} \right] dM_D, \\ P_R(M_R) dM_R = \left(\frac{1}{\pi\Lambda_R^2} \right)^{\beta N(N-1)/4} \left(\frac{1}{2\pi\Lambda_R^2} \right)^{\beta N/2} \\ \times \exp \left[-\frac{\text{tr}(M_R^\dagger M_R)}{2\Lambda_R^2} \right] dM_R, \quad (3.1)$$

where $\beta = 1$ corresponds to real matrix elements and $\beta = 2$ corresponds to complex matrix elements. The difference between the two PDFs comes from the fact that M_R is symmetric while M_D is not. The PDFs (3.1) are properly normalized with respective standard deviations Λ_D and Λ_R (for the diagonal elements only, the off-diagonal elements have standard deviations $\Lambda_R/\sqrt{2}$).

Before proceeding, it is convenient to reexpress the PDFs in terms of the dimensionless variables $\hat{M}_D = M_D/(\sqrt{2}\Lambda_D)$ and $\hat{M}_R = M_R/(\sqrt{2}\Lambda_R)$. Hence the PDFs (3.1) become

$$P_D(\hat{M}_D) d\hat{M}_D = \tilde{C}_{DN}^\beta e^{-\text{tr}(\hat{M}_D^\dagger \hat{M}_D)} d\hat{M}_D, \\ P_R(\hat{M}_R) d\hat{M}_R = \tilde{C}_{RN}^\beta e^{-\text{tr}(\hat{M}_R^\dagger \hat{M}_R)} d\hat{M}_R, \quad (3.2)$$

where the normalization constants are

$$\tilde{C}_{DN}^\beta = \frac{1}{\pi^{\beta N^2/2}} \quad \text{and} \quad \tilde{C}_{RN}^\beta = \frac{2^{\beta N(N-1)/4}}{\pi^{\beta N(N+1)/4}}. \quad (3.3)$$

By the singular (eigen) value decomposition theorem, both matrices can be diagonalized with non-negative (real) elements as $\hat{M}_D = U_L D_D U_R^\dagger$ and $\hat{M}_R = U D_R U^T$ where the U -matrices are unitary (orthogonal) for $\beta = 2$ ($\beta = 1$). In terms of the singular values and the remaining variables, the measure for the Dirac neutrino mass matrix can be simplified to

$$d\hat{M}_D = c_{AN}^\beta \prod_{1 \leq i < j \leq N} |\hat{m}_{D,i}^2 - \hat{m}_{D,j}^2|^\beta \\ \times \prod_{1 \leq i \leq N} \hat{m}_{D,i}^{\beta-1} d\hat{m}_{D,i} U_{LR}^\dagger dU_{LR}, \quad (3.4)$$

where the measure $d\hat{M}'_D = d\hat{M}_D$ is invariant under a transformation $\hat{M}'_D = U_L \hat{M}_D U_R^\dagger$ and $U_{LR}^\dagger dU_{LR}$ is the appropriate group (Haar) measure. Moreover, the measure for the right-handed neutrino mass matrix, which is invariant under $\hat{M}'_R = U \hat{M}_R U^T$, i.e. $d\hat{M}'_R = d\hat{M}_R$, can be written as

$$d\hat{M}_R = c_{SN}^\beta \prod_{1 \leq i < j \leq N} |\hat{m}_{R,i}^\beta - \hat{m}_{R,j}^\beta| \prod_{1 \leq i \leq N} |\hat{m}_{R,i}|^{\beta-1} d\hat{m}_{R,i} U^\dagger dU, \quad (3.5)$$

where $U^\dagger dU$ is again the appropriate group (Haar) measure. The proofs of (3.4) and (3.5) are given in Appendix A. The normalization constants for arbitrary and symmetric matrices c_{AN}^β and c_{SN}^β can also be found in the Appendix and are given by (A9) and (A5) respectively. The absolute value is not necessary for the Dirac matrix because all singular values are non-negative. On the other hand, the absolute value is absolutely necessary for the real Majorana matrix ($\beta = 1$) because the decomposition in this case is an eigenvalue decomposition where all eigenvalues can also be negative. In the following, the expression ‘‘singular values’’ will be used for both singular values and eigenvalues.

Since $P_D(\hat{M}_D)$ and $P_R(\hat{M}_R)$ depend only on the singular values, it is easy to write the joint PDFs for the singular values,

$$\begin{aligned} P_D(\hat{m}_D) d\hat{m}_D &= C_{DN}^\beta \prod_{1 \leq i < j \leq N} |\hat{m}_{D,i}^2 - \hat{m}_{D,j}^2|^\beta \\ &\quad \times \prod_{1 \leq i \leq N} \hat{m}_{D,i}^{\beta-1} e^{-\hat{m}_{D,i}^2} d\hat{m}_{D,i}, \\ P_R(\hat{m}_R) d\hat{m}_R &= C_{RN}^\beta \prod_{1 \leq i < j \leq N} |\hat{m}_{R,i}^\beta - \hat{m}_{R,j}^\beta| \\ &\quad \times \prod_{1 \leq i \leq N} |\hat{m}_{R,i}|^{\beta-1} e^{-\hat{m}_{R,i}^2} d\hat{m}_{R,i}, \end{aligned} \quad (3.6)$$

where the normalization constants are

$$\begin{aligned} C_{DN}^\beta &= \tilde{C}_{DN}^\beta c_{AN}^\beta \int U_{LR}^\dagger dU_{LR} \\ &= \tilde{C}_{DN}^\beta c_{AN}^\beta \frac{[\text{Vol}(\mathcal{V}_N^\beta)]^2}{(2\pi)^{(\beta-1)N}} \\ &= \frac{2^N \pi^{N(1-\beta/2)}}{N!} \prod_{1 \leq i \leq N} \frac{1}{[\Gamma(\beta i/2)]^2}, \\ C_{RN}^\beta &= \tilde{C}_{RN}^\beta c_{SN}^\beta \int U^\dagger dU \\ &= \tilde{C}_{RN}^\beta c_{SN}^\beta \text{Vol}(\mathcal{V}_N^\beta) \\ &= \frac{2^{N[\beta(N+3)-4]/4}}{N!} \prod_{1 \leq i \leq N} \frac{1}{\Gamma(\beta i/2)}. \end{aligned} \quad (3.7)$$

The volume of the Stiefel manifold $\mathcal{V}_{N,N}^\beta \equiv \mathcal{V}_N^\beta$ is given in (A1) [12]. Note that the normalization constants can also be obtained from well-known Selberg-like integrals [7]. For example, one has

$$C_{DN}^\beta = 2^N \prod_{1 \leq i \leq N} \frac{\Gamma(\beta/2 + 1)}{\Gamma(\beta i/2 + 1) \Gamma(\beta i/2)},$$

for the Dirac case. This identity can be obtained by exploiting the properties of the function $\Gamma(x)$.

B. Jacobian

In terms of the dimensionless Dirac and Majorana neutrino mass matrices \hat{M}_D and \hat{M}_R , the light neutrino mass matrix is given by

$$M_\nu = \sqrt{2} \Lambda_\nu \hat{M}_\nu = -\frac{\sqrt{2} \Lambda_D^2}{\Lambda_R} \hat{M}_D \hat{M}_R^{-1} \hat{M}_D^T,$$

where \hat{M}_ν is the dimensionless light neutrino mass matrix and $\Lambda_\nu = \Lambda_D^2/\Lambda_R$ is the light neutrino mass scale. Hence $\hat{M}_\nu = -\hat{M}_D \hat{M}_R^{-1} \hat{M}_D^T$. To determine the PDF for \hat{M}_ν from the PDFs of \hat{M}_D and \hat{M}_R , it is necessary to compute the norm of the determinant of the Jacobian matrix corresponding to the appropriate change of variables. Since \hat{M}_R and \hat{M}_ν have the same number of independent parameters, this computation simplifies greatly when the change of variables is chosen from (\hat{M}_D, \hat{M}_R) to (\hat{M}_D, \hat{M}_ν) . Thus the Jacobian matrix J^β is given schematically by

$$J^\beta = \left[\frac{\partial \hat{M}_R}{\partial \hat{M}_\nu} \right] = - \left[\frac{\partial (\hat{M}_D^T \hat{M}_\nu^{-1} \hat{M}_D)}{\partial \hat{M}_\nu} \right]. \quad (3.8)$$

This Jacobian can be easily obtained with the help of the wedge product technique [8]. Indeed, since for a symmetric matrix $X = CYC^T$ where C is a constant nonsingular matrix, the measure satisfies $dX = CdYC^T$ and the wedge product leads to

$$(dX) \equiv \bigwedge_{1 \leq i < j \leq N} dX_{ij} = (CdYC^T) = p(C)(dY),$$

where $p(C)$ is a polynomial in C . Using $C = C_2 C_1$, one obtains that $p(C_2 C_1) = p(C_1) p(C_2)$. The only function satisfying this condition is a positive power of the determinant, hence $p(C) = (\det C)^k$ with $k \geq 0$. To obtain k , it suffices to choose a simple matrix, for example $C = \text{diag}(c, 1, \dots, 1)$ for which

$$X = \begin{pmatrix} c^2 y_{11} & c y_{12} & \cdots & c y_{1n} \\ c y_{12} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c y_{1n} & y_{2n} & \cdots & y_{nn} \end{pmatrix}.$$

Therefore, the wedge product for a real matrix leads to

$$\begin{aligned} (dX) &= (CdYC^T) \equiv \bigwedge_{1 \leq i < j \leq N} (CdYC^T)_{ij} \\ &= c^{N+1} \bigwedge_{1 \leq i < j \leq N} dY_{ij} = c^{N+1} (dY), \end{aligned}$$

while for a complex matrix the result becomes

$$\begin{aligned}
 (dX) &= (CdYC^T) \equiv \bigwedge_{1 \leq i \leq j \leq N} \text{Re}(CdYC^T)_{ij} \wedge \text{Im}(CdYC^T)_{ij} \\
 &= c^{2(N+1)} \bigwedge_{1 \leq i \leq j \leq N} d\text{Re}Y_{ij} \wedge d\text{Im}Y_{ij} = c^{2(N+1)}(dY).
 \end{aligned}$$

Thus one concludes that the measure satisfies

$$dX = |\det C|^{\beta(N+1)} dY, \quad (3.9)$$

i.e. the positive power is $k = \beta(N+1)$.

It is now straightforward to obtain the Jacobian of interest. Indeed, using (3.9) one has

$$\begin{aligned}
 d\hat{M}_R &= -\hat{M}_D^T d\hat{M}_\nu^{-1} \hat{M}_D = \hat{M}_D^T \hat{M}_\nu^{-1} d\hat{M}_\nu \hat{M}_\nu^{-1} \hat{M}_D \\
 &= |\det(\hat{M}_\nu^{-1} \hat{M}_D)|^{\beta(N+1)} d\hat{M}_\nu
 \end{aligned}$$

where the identity $d\hat{M}_\nu^{-1} = -\hat{M}_\nu^{-1} d\hat{M}_\nu \hat{M}_\nu^{-1}$ was used.

Finally, the full measure is simply

$$d\hat{M}_D d\hat{M}_R = \left| \frac{\det \hat{M}_D}{\det \hat{M}_\nu} \right|^{\beta(N+1)} d\hat{M}_D d\hat{M}_\nu. \quad (3.10)$$

It is now straightforward to verify that the measure $d\hat{M}_\nu$ is invariant under every transformation $\hat{M}_\nu \rightarrow \hat{M}'_\nu = U_\nu \hat{M}_\nu U_\nu^T$ where $U_\nu \in \{O(N), U(N)\}$ for $\beta = \{1, 2\}$ respectively. Indeed, since $d\hat{M}'_D = d\hat{M}_D$ and $d\hat{M}'_R = d\hat{M}_R$ for $\hat{M}'_D = U_L \hat{M}_D U_R^\dagger$ and $\hat{M}'_R = U \hat{M}_R U^T$ respectively, one obtains

$$\begin{aligned}
 0 &= d\hat{M}'_D d\hat{M}'_R - d\hat{M}_D d\hat{M}_R \\
 &= \left| \frac{\det \hat{M}'_D}{\det \hat{M}'_\nu} \right|^{\beta(N+1)} d\hat{M}'_D d\hat{M}'_\nu - \left| \frac{\det \hat{M}_D}{\det \hat{M}_\nu} \right|^{\beta(N+1)} d\hat{M}_D d\hat{M}_\nu \\
 &= \left| \frac{\det \hat{M}_D}{\det \hat{M}_\nu} \right|^{\beta(N+1)} d\hat{M}_D (d\hat{M}'_\nu - d\hat{M}_\nu),
 \end{aligned}$$

with $\hat{M}'_\nu = -\hat{M}'_D \hat{M}'_R^{-1} \hat{M}'_D^T = U_L \hat{M}_\nu U_L^T$ and $U = U_R^*$. Therefore, the measure $d\hat{M}'_\nu = d\hat{M}_\nu$ for every transformation in the appropriate group as stated above. Hence $d\hat{M}_\nu$ can be written in terms of its singular values in exactly the same way as $d\hat{M}_R$ in (3.5).

C. Probability density function

Now that the norm of the determinant of the Jacobian matrix (3.8) has been found, it is simple to determine the PDF $P_\nu(\hat{M}_\nu) d\hat{M}_\nu$. Indeed, with the help of the change of variables from (\hat{M}_D, \hat{M}_R) to (\hat{M}_D, \hat{M}_ν) , the PDF is calculated from $P_D(\hat{M}_D) P_R(\hat{M}_R) d\hat{M}_D d\hat{M}_R = P_{D\nu}(\hat{M}_D, \hat{M}_\nu) d\hat{M}_D d\hat{M}_\nu$ and the measure (3.10), which leads to

$$\begin{aligned}
 P_{D\nu}(\hat{M}_D, \hat{M}_\nu) d\hat{M}_D d\hat{M}_\nu &= P_D(\hat{M}_D) P_R(-\hat{M}_D^T \hat{M}_\nu^{-1} \hat{M}_D) \\
 &\quad \times |\det J^\beta| d\hat{M}_D d\hat{M}_\nu. \quad (3.11)
 \end{aligned}$$

The PDF $P_\nu(\hat{M}_\nu) d\hat{M}_\nu$ is obtained from (3.11) after marginalizing over the variables \hat{M}_D , which gives

$$\begin{aligned}
 P_\nu(\hat{M}_\nu) d\hat{M}_\nu &= \tilde{C}_{\nu N}^\beta \left[\int e^{-\text{tr}(\hat{M}_D^\dagger \hat{M}_D) - \text{tr}(\hat{M}_D^\dagger \hat{M}_\nu^{-1} \hat{M}_D^* \hat{M}_\nu^T \hat{M}_\nu^{-1} \hat{M}_D)} \right. \\
 &\quad \left. \times \left| \frac{\det \hat{M}_D}{\det \hat{M}_\nu} \right|^{\beta(N+1)} d\hat{M}_D \right] d\hat{M}_\nu, \quad (3.12)
 \end{aligned}$$

where (3.3) give the normalization constant as

$$\tilde{C}_{\nu N}^\beta = \tilde{C}_{DN}^\beta \tilde{C}_{RN}^\beta = \frac{2^{\beta N(N-1)/4}}{\pi^{\beta N(3N+1)/4}}. \quad (3.13)$$

In terms of the light neutrino mass matrix singular values $\hat{m}_{\nu,i}$ and group variables U_ν , the neutrino PDF (3.12) becomes

$$\begin{aligned}
 P_\nu(\hat{m}_\nu) d\hat{m}_\nu \frac{U_\nu^\dagger dU_\nu}{\text{Vol}(\mathcal{V}_N^\beta)} &= \left[\int P_D(\hat{M}_D) P_R(-\hat{M}_D^T D_\nu^{-1} \hat{M}'_D) |\det J^\beta| d\hat{M}_D \right] c_{SN}^\beta \prod_{1 \leq i < j \leq N} |\hat{m}_{\nu,i}^\beta - \hat{m}_{\nu,j}^\beta| \prod_{1 \leq i \leq N} |\hat{m}_{\nu,i}|^{\beta-1} d\hat{m}_{\nu,i} \frac{U_\nu^\dagger dU_\nu}{\text{Vol}(\mathcal{V}_N^\beta)} \\
 &= C_{\nu N}^\beta I_N^\beta(\hat{m}_{\nu,1}, \dots, \hat{m}_{\nu,N}) \prod_{1 \leq i < j \leq N} |\hat{m}_{\nu,i}^\beta - \hat{m}_{\nu,j}^\beta| \prod_{1 \leq i \leq N} |\hat{m}_{\nu,i}|^{-(\beta N+1)} d\hat{m}_{\nu,i} \frac{U_\nu^\dagger dU_\nu}{\text{Vol}(\mathcal{V}_N^\beta)}, \quad (3.14)
 \end{aligned}$$

where the normalization constant is obtained from (3.13) and is given by

$$C_{\nu N}^\beta = \frac{\tilde{C}_{\nu N}^\beta c_{SN}^\beta c_{AN}^\beta [\text{Vol}(\mathcal{V}_N^\beta)]^3}{2^N (2\pi)^{(\beta-1)N}} = \frac{2^{N[\beta(N+3)-4]/4}}{N!} \prod_{1 \leq i \leq N} \frac{\Gamma(\beta/2 + 1)}{\Gamma(\beta i/2 + 1) [\Gamma(\beta i/2)]^2}, \quad (3.15)$$

while the remaining function is

$$\begin{aligned}
 I_N^\beta(t_1, \dots, t_N) &= \frac{2^N}{\text{Vol}(\mathcal{V}_N^\beta)} \int_{U_L, U_R \in \mathcal{V}_N^\beta} \int_0^\infty e^{-\sum_{1 \leq i, j \leq N} |\sum_{1 \leq k \leq N} t_k^{-1} (U_L)_{ki} (U_L)_{kj}|^2 \hat{m}_{D,i}^2 \hat{m}_{D,j}^2} \\
 &\quad \times \prod_{1 \leq i < j \leq N} |\hat{m}_{D,i}^2 - \hat{m}_{D,j}^2|^\beta \prod_{1 \leq i \leq N} e^{-\hat{m}_{D,i}^2} |\hat{m}_{D,i}|^{\beta(N+2)-1} d\hat{m}_{D,i} \frac{(U_L^\dagger dU_L)' (U_R^\dagger dU_R)'}{\text{Vol}(\mathcal{V}_N^\beta) / (2\pi)^{(\beta-1)N}} \\
 &= \int_{U \in \mathcal{V}_N^\beta} \int_0^\infty \prod_{1 \leq i < j \leq N} |x_i - x_j|^\beta e^{-2|\sum_{1 \leq k \leq N} t_k^{-1} U_{ki} U_{kj}|^2 x_i x_j} \\
 &\quad \times \prod_{1 \leq i \leq N} x_i^{\beta(N+2)/2-1} e^{-x_i(1+|\sum_{1 \leq j \leq N} t_j^{-1} U_{ji}^2| x_i)} dx_i \frac{(U^\dagger dU)'}{\text{Vol}(\mathcal{V}_N^\beta) / (2\pi)^{(\beta-1)N}}. \tag{3.16}
 \end{aligned}$$

In deriving (3.14), the change of variables $\hat{M}_\nu = U_\nu D_\nu U_\nu^T$ where $D_\nu = \text{diag}(\hat{m}_{\nu,1}, \dots, \hat{m}_{\nu,N})$ was used. The invariance of the measure $d\hat{M}_D$ under the appropriate transformations was also necessary. To rewrite the function I_N^β , the Dirac neutrino mass matrix elements was expressed in terms of its singular values and its group variables. The group variables for U_R were integrated over straightforwardly while the group variables for U_L (minus the phases, hence the prime) remain due to their complicated coupling with the singular values. Finally, the change of variables $x_i = \hat{m}_{D,i}^2$ and $U_L = U$ were done to simplify the integral I_N^β .

Note that integration over the light neutrino group variables U_ν in (3.14) is straightforward, as in (3.6), leading to

$$\begin{aligned}
 P_\nu(\hat{m}_\nu) d\hat{m}_\nu &= C_{\nu N}^\beta I_N^\beta(\hat{m}_{\nu,1}, \dots, \hat{m}_{\nu,N}) \prod_{1 \leq i < j \leq N} |\hat{m}_{\nu,i}^\beta - \hat{m}_{\nu,j}^\beta| \\
 &\quad \times \prod_{1 \leq i \leq N} |\hat{m}_{\nu,i}|^{-(\beta N+1)} d\hat{m}_{\nu,i}, \tag{3.17}
 \end{aligned}$$

for the singular value PDF, with the normalization constant (3.15). From (3.14), the PDF for the group variables U_ν is thus uniquely determined by the appropriate group (Haar) measure (minus the phases). This behavior is problematic since there must be a correlation between the neutrino masses (i.e. the singular values of \hat{M}_ν) and the neutrino mass eigenstates (i.e. the singular vectors of \hat{M}_ν which are given by the columns of the neutrino mixing matrix U_ν). For example, the mostly electronic neutrino must be the lightest (normal hierarchy) or the second lightest (inverted hierarchy). This simple observation has important consequences in the analysis of the physical case appropriate for the SM.

For general N , the function (3.16) makes the analysis of the seesaw ensemble PDF (3.17) quite intricate since it is not only a random matrix theory PDF, it is a random matrix theory PDF that cannot be written as an analytic function (apart for the $N = 1$ case, to the best of our knowledge).

For real matrices, none of the group variables in (3.16) can be straightforwardly integrated. For complex matrices however, the integration over group variables does not

include some of the phases, hence the prime as explained in Appendix A. This can also be seen by using an adequate parametrization of the matrix U in (3.16). Indeed, following in part [13], it is possible to parametrize any $N \times N$ unitary matrix as

$$U = \prod_{1 \leq j < k \leq N} \exp(i\phi_{jk} P_k) \exp(i\theta_{jk} \Sigma_{jk}) \prod_{1 \leq j \leq N} \exp(i\varphi_j P_j),$$

where the matrices P_j and Σ_{jk} are given by

$$P_{j,ik} = \delta_{ji} \delta_{jk}, \quad \Sigma_{jk,i\ell} = -i\delta_{ji} \delta_{k\ell} + i\delta_{j\ell} \delta_{ki},$$

and the range of the N^2 mixing angles θ_{jk} and phases ϕ_{jk} and φ_j are

$$\theta_{jk} \in [0, \pi/2), \quad \phi_{jk} \in [0, 2\pi), \quad \varphi_j \in [0, 2\pi).$$

With this parametrization it is clear that the integrand in (3.16) is independent of the N phases φ_j , explaining why they are not integrated over in (3.16).

IV. PROPERTIES OF THE SEESAW ENSEMBLE

This section discusses the general properties of the seesaw ensemble PDF (3.17). The asymptotic behaviors at $\hat{m}_{\nu,i} \rightarrow 0$ and $\hat{m}_{\nu,i} \rightarrow \infty$ are investigated. Moreover, the $N = 1$ case as well as the large N case are studied analytically.

A. Asymptotic Behaviors

First, being a PDF, (3.17) implies the following identity for $\beta = 1$ or $\beta = 2$:

$$\begin{aligned}
 \int_{\hat{m}_{\nu,\min}}^\infty I_N^\beta(\hat{m}_{\nu,1}, \dots, \hat{m}_{\nu,N}) \prod_{1 \leq i < j \leq N} |\hat{m}_{\nu,i}^\beta - \hat{m}_{\nu,j}^\beta| \\
 \times \prod_{1 \leq i \leq N} |\hat{m}_{\nu,i}|^{-(\beta N+1)} d\hat{m}_{\nu,i} &= \frac{1}{C_{\nu N}^\beta},
 \end{aligned}$$

where $\hat{m}_{\nu,\min} = -\infty$ for $\beta = 1$ (eigenvalues) and $\hat{m}_{\nu,\min} = 0$ for $\beta = 2$ (singular values).

The PDF (3.17) is difficult to study analytically for general N since the function (3.16) is hard to evaluate generally. It is nevertheless possible to investigate the asymptotic behaviors of the PDF (3.17) for general N as follows.

First, as $\hat{m}_{\nu,i} \rightarrow \pm\infty$ for a fixed i , one has

$$\begin{aligned} P_\nu(\hat{m}_\nu) &= k_\infty |\hat{m}_{\nu,i}|^{\beta(N-1)-(\beta N+1)} [1 + \mathcal{O}(|\hat{m}_{\nu,i}|^{-1})] \\ &= k_\infty |\hat{m}_{\nu,i}|^{-(\beta+1)} [1 + \mathcal{O}(|\hat{m}_{\nu,i}|^{-1})]. \end{aligned}$$

$$\begin{aligned} I_N^\beta(\hat{m}_{\nu,1}, \dots, \hat{m}_{\nu,N}) &\sim |\hat{m}_{\nu,i}|^{\beta N(2N+1)/2} \int_{U \in \mathcal{U}_N^\beta} \int_0^\infty \prod_{1 \leq j < k \leq N} |y_j - y_k|^\beta e^{-2\hat{m}_{\nu,i}^2 \sum_{1 \leq \ell \leq N} \hat{m}_{\nu,\ell}^{-1} U_{\ell j} U_{\ell k} |y_j y_k|} \\ &\quad \times \prod_{1 \leq j \leq N} y_j^{\beta(N+2)/2-1} e^{-|\hat{m}_{\nu,i}| y_j (1+|\hat{m}_{\nu,i}| \sum_{1 \leq k \leq N} \hat{m}_{\nu,k}^{-1} U_{kj}^2 |y_j|)} dy_j (U^\dagger dU)', \end{aligned}$$

and that implies that for $\hat{m}_{\nu,i} \rightarrow 0$, the PDF behaves as

$$P_\nu(\hat{m}_\nu) = k_0 |\hat{m}_{\nu,i}|^{\beta N(2N-1)/2-1} [1 + \mathcal{O}(|\hat{m}_{\nu,i}|)].$$

Therefore the PDF vanishes at $\hat{m}_{\nu,i} = 0$ unless $N = 1$. Again the constant k_0 is a function of the remaining singular values and N that is hard to evaluate exactly.

With the help of both asymptotic behaviors, it is straightforward to conclude that the most probable singular value is a finite number different than zero, unless $N = 1$. In other words, for $N \geq 2$ the seesaw ensemble prefers the lightest neutrino to be massive.

B. The $N = 1$ case

For now, the only analytic case with finite N corresponds to $N = 1$, for which the joint PDF is given by

Therefore, as long as the asymptotic expansion converges uniformly, the average singular value (and all moments greater than 1) is not well defined for $\beta = 1$ while the standard deviation (and all moments greater than 2) is not well defined for $\beta = 2$. Here the constant k_∞ depends on the other singular values and on N and is thus difficult to calculate analytically.

The limiting behavior at $\hat{m}_{\nu,i} \rightarrow 0$ is obtained from the function I_N^β . From (3.16) and the rescaling $x_j = |\hat{m}_{\nu,i}| y_j$, one gets

$$\begin{aligned} P_\nu(\hat{m}_\nu) d\hat{m}_\nu &= \frac{\pi^{1-\beta/2} \Gamma(3\beta/2)}{2^{\beta/2+1} [\Gamma(\beta/2)]^3} |\hat{m}_\nu|^{(\beta-2)/2} \\ &\quad \times U(3\beta/4, 1/2, \hat{m}_\nu^2/4), \end{aligned} \quad (4.1)$$

where $U(a, b, z)$ is the confluent hypergeometric function of the second kind. The joint PDF (4.1) can be obtained analytically because the function (3.16) is simple in the $N = 1$ case. The plot of (4.1) is shown in Fig. 1 along with numerical results. The agreement between the two approaches is clear.

Figure 1 shows that the most probable value for the singular value is $\hat{m}_\nu = 0$. This is easily verified by looking at $\partial P_\nu(\hat{m}_\nu)/\partial \hat{m}_\nu$ which is always negative (positive) for positive (negative) \hat{m}_ν . Moreover, for $\beta = 2$, the PDF of the light neutrino phase is the Haar measure which is flat and therefore uninteresting. Also, from the limiting behaviors, which can be obtained directly from

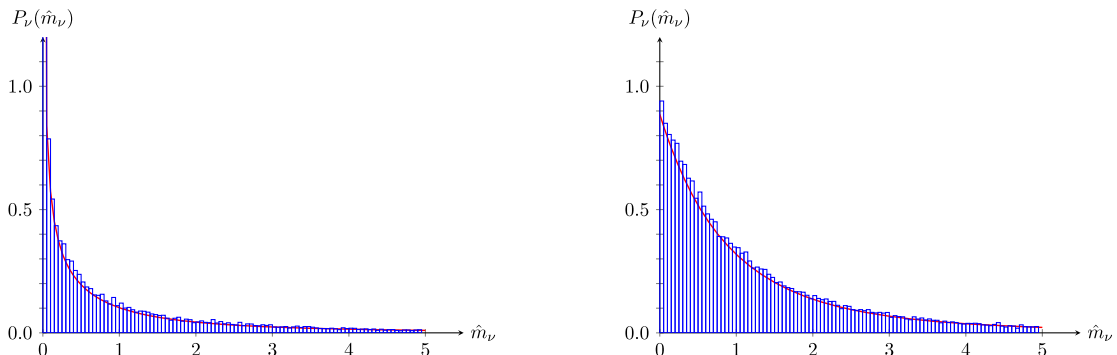


FIG. 1. Probability density function for the seesaw ensemble with $N = 1$. The red curve corresponds to the analytic result while the histogram corresponds to numerical results (with 5×10^4 dimensionless light neutrino mass matrices generated). The left panel shows matrices with real elements ($\beta = 1$) while the right panel shows matrices with complex elements ($\beta = 2$).

$$U(3\beta/4, 1/2, z) = \begin{cases} z^{-3\beta/4} [1 + \mathcal{O}(z^{-1})] & \text{for } z \rightarrow \infty \\ \frac{\pi^{1/2}}{\Gamma[(3\beta+2)/4]} [1 + \mathcal{O}(z^{1/2})] & \text{for } z \rightarrow 0 \end{cases},$$

none of the moments exist except for the average value of the singular value for $\beta = 2$. A simple computation using the following Mellin transforms:

$$\int_0^\infty z^{\lambda-1} U(a, b, z) dz = \frac{\Gamma(\lambda)\Gamma(a-\lambda)\Gamma(\lambda-b+1)}{\Gamma(a)\Gamma(a-b+1)} \\ \text{for } \max(b-1, 0) < \lambda < a,$$

shows it is given by $\langle \hat{m}_\nu \rangle_{N=1}^{\beta=2} = \sqrt{\pi}$. Since (4.1) is an even function of \hat{m}_ν , one can nevertheless formally define the average value for $\beta = 1$, and it is $\langle \hat{m}_\nu \rangle_{N=1}^{\beta=1} = 0$. These observations agree with the general analysis of Sec. IV A.

C. The large N case

For usual Gaussian ensembles, both the $N = 1$ PDFs and the level densities at large N are simple. Indeed, for the usual Gaussian random matrix theory, one can show the celebrated Wigner's semicircle law using the moment method or the resolvent method. It is therefore plausible that the level density at large N for the seesaw ensemble is also a simple analytic function. Moreover, from the $1/N$ expansion, the large N case could shed some light on the physical case of $N = 3$ neutrinos, as occurs for example with quantum chromodynamics. It is however impossible here to use the moment method or the resolvent method since both methods rely on the computation of the moments and most moments do not exist for the seesaw ensemble, as argued above.

Another option is to translate the problem to a Coulomb-like gas and take the thermodynamic limit, which effectively corresponds to the limit $N \rightarrow \infty$. To do so, the singular value PDF (3.17) is reexpressed as a thermal system $e^{-H_{\nu N}^\beta}$ where the temperature is fixed. In this picture, the Hamiltonian for the seesaw ensemble is obtained from (3.17) and is given by

$$H_{\nu N}^\beta = - \sum_{1 \leq i < j \leq N} \ln |x_i^\beta - x_j^\beta| + (\beta N + 1) \sum_{1 \leq i \leq N} \ln |x_i| \\ - \ln I_N^\beta(x_1, \dots, x_N), \quad (4.2)$$

where the positions of the N charged particles are given by x_i . It is important to note that the Hamiltonian (4.2) corresponds to a Coulomb gas only for $\beta = 1$. Indeed, the characteristic logarithmic two-dimensional Coulomb potential occurs only for $\beta = 1$.

The thermodynamic limit is computable with the help of the level density of singular values $\rho_{\nu N}^\beta(x)$, which is defined as the following correlation function:

$$\rho_{\nu N}^\beta(x) = N \int P_\nu(x, \hat{m}_{\nu,2}, \dots, \hat{m}_{\nu,N}) \prod_{2 \leq i \leq N} d\hat{m}_{\nu,i},$$

such that $\int \rho_{\nu N}^\beta(x) dx = N$ the number of charged particles. Assuming the distance between charged particles decreases as N tends to infinity, the Hamiltonian (4.2) in the large N limit can be approximated by a continuum of singular values with Hamiltonian,

$$H_{\nu N}^\beta = -\frac{1}{2} \int dx dy \rho_{\nu N}^\beta(x) \rho_{\nu N}^\beta(y) \ln |x^\beta - y^\beta| \\ + \int dx \rho_{\nu N}^\beta(x) V_{\nu N}^\beta, \\ V_{\nu N}^\beta = (\beta N + 1) \ln |x| + \dots, \quad (4.3)$$

where the ellipses represents the contribution from the integral I_N^β to the potential $V_{\nu N}^\beta$. Extremizing the Hamiltonian (4.3) with respect to the level density of singular values subject to the constraint $\int \rho_{\nu N}^\beta(x) dx = N$ leads to an equation for the level density of singular values which can usually be solved analytically. This is done in Appendix B for real and complex arbitrary matrices as well as real and complex symmetric matrices with Gaussian ensembles. This is however not the case here due to the complexity of the integral I_N^β . Indeed, for usual Gaussian ensembles, the potential $V_{\nu N}^\beta \propto x^2$ depends only on one variable and extremization of the Hamiltonian is straightforward (see Appendix B). For the seesaw ensemble however, the potential $V_{\nu N}^\beta$ is a complicated integral function of several variables, thus extremizing the Hamiltonian is not so simple. Moreover, since e^{-1/x^2} does not have an expansion around $x = 0$, it is not possible to compute the first contributions to the potential from the integral I_N^β .

To proceed, it is proposed to approximate the problem by computing the level densities in the thermodynamic limit for both the Dirac and Majorana neutrino mass matrices first, and then marginalizing to get the level density of singular values for the seesaw ensemble. Since interchanging the order of the steps gets rid of all complications related to the group variables, this technique can at best give an approximation to the true level density of singular values. Although a comparison between the analytical approximation obtained with this technique and numerical results shows that the approximation is quite good, it will be argued that the results are wrong.

The level densities for the Dirac (arbitrary) and Majorana (symmetric) matrices are given in (B5). Following the same procedure as in Sec. III C and marginalizing with the approximation $\hat{x}_R = \hat{x}_D^2 / \hat{x}$ where $\hat{x} \equiv \hat{x}_\nu$, which is the analog of $\hat{M}_R = -\hat{M}_D^T \hat{M}_\nu^{-1} \hat{M}_D$, gives

$$\hat{\rho}_{\nu N}^\beta(\hat{x}) = \int d\hat{x}_D \hat{\rho}_D(\hat{x}_D) \hat{\rho}_R(\hat{x}_D^2 / \hat{x}) |\det \hat{J}|.$$

Here the normalized quantities $\hat{x} = x / \sqrt{N}$ and $\hat{\rho}(\hat{x}) = \rho(x) / \sqrt{N}$ are introduced in Appendix B. Moreover, the

change of variables leads to the Jacobian $|\det \hat{J}| = \hat{x}_D^2 / \hat{x}^2$. The solution to the marginalization procedure of the two level densities is given by

$$\hat{\rho}_{\nu N}^\beta(\hat{x}) = \begin{cases} \frac{\beta^{3/4}}{10\pi^{5/2}|\hat{x}|^{1/2}} \left[\Gamma\left(-\frac{1}{4}\right)^2 {}_3F_2\left(-\frac{1}{4}, \frac{1}{4}, \frac{3}{4}; \frac{1}{2}, \frac{9}{4}; \frac{\hat{x}^2}{4\beta}\right) - \frac{10}{21} \frac{|\hat{x}|}{2\beta^{1/2}} \Gamma\left(\frac{1}{4}\right)^2 {}_3F_2\left(\frac{1}{4}, \frac{3}{4}, \frac{5}{4}; \frac{3}{2}, \frac{11}{4}; \frac{\hat{x}^2}{4\beta}\right) \right] & \text{for } |\hat{x}| < 2\sqrt{\beta}. \\ \frac{\beta^{3/2}}{\pi\hat{x}^2} {}_3F_2\left(-\frac{1}{2}, \frac{3}{4}, \frac{5}{4}; \frac{3}{2}, 2; \frac{4\beta}{\hat{x}^2}\right) & \text{for } |\hat{x}| \geq 2\sqrt{\beta} \end{cases} \quad (4.4)$$

The continuity at $\hat{x} = 2\sqrt{\beta}$ and the normalization $\int d\hat{x} \hat{\rho}_{\nu N}^\beta(\hat{x}) = 1$ of the level density (4.4) can be verified exactly by using generalizations of the Gauss's identity ${}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$ for $c - a - b > 0$. The level density for the seesaw ensemble (4.4) is shown in Fig. 2 with a comparison to numerical results. The behaviors of the level density around zero and infinity (more precisely $|\hat{x}| \gtrsim 2$) match the numerical results quite well while the intermediate regime is not as good. The agreement is nonetheless satisfying considering the approximations made.

The limiting behaviors at $\hat{x} \rightarrow 0$ and $\hat{x} \rightarrow \pm\infty$ are

$$\hat{\rho}_{\nu N}^\beta(\hat{x}) = \begin{cases} \frac{\beta^{3/4}\Gamma(-\frac{1}{4})^2}{10\pi^{5/2}|\hat{x}|^{1/2}} & \text{for } |\hat{x}| \rightarrow 0 \\ \frac{\beta^{3/2}}{\pi\hat{x}^2} & \text{for } |\hat{x}| \rightarrow \infty \end{cases},$$

and they imply that none of the moments exist although the level density is integrable. This is consistent with the results obtained in Sec. IV A from the full PDF for $\beta = 1$ but not for $\beta = 2$. It is therefore expected that the match between the approximated level density (4.4) and numerical results is better for $\beta = 1$ compared to $\beta = 2$. Although both the real and complex cases seem to agree with the numerical results as seen in Fig. 2, it is clear from the difference in the moments that the level density at large N (4.4) is wrong for $\beta = 2$. The error originates from all the simplifications

made, most likely from the Jacobian of the transformation. Indeed, the expected β dependence of the Jacobian, as in (3.10), does not occur due to the approximations, making the analogy with the full PDF possibly correct only for $\beta = 1$. Hence computing the level densities at large N for the Dirac and Majorana neutrino mass matrices and then marginalizing to obtain the level density at large N for the seesaw ensemble do not commute for $\beta = 2$.

Furthermore, contrary to usual level densities, the level density for the seesaw ensemble does not have compact support due to the inverse Majorana neutrino mass matrix that appears in the marginalization. Hence there are arbitrarily large singular values, although the level density and the mean level spacing, given by $1/\hat{\rho}_{\nu N}^\beta(\hat{x})$, show that large singular values are rare and far apart from each other.

Finally, apart from the fact that the approximated analytic level density seems wrong, it seems ill advised to work back and compute the potential for the seesaw ensemble from the level density since it is already known that $V_{\nu N}^\beta$ depends on more than one variable.

D. Comparison between $N = 1$ and large N

Before concluding, it is interesting to compare the results for $N = 1$ and large N . Indeed, although those regimes are as far apart as possible, by definition the PDF for $N = 1$ is the $N = 1$ (normalized) level density. Hence, (4.1) can be compared directly with (4.4), allowing a qualitative

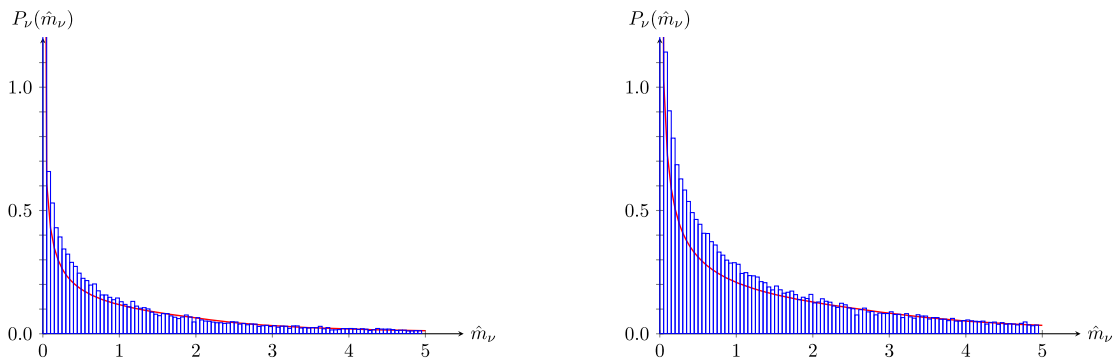


FIG. 2. Level density at large N for the seesaw ensemble. The red curve corresponds to the approximated analytic result while the histogram corresponds to numerical results (with 10^3 singular values generated from 50×50 matrices). The left panel shows matrices with real elements ($\beta = 1$) while the right panel shows matrices with complex elements ($\beta = 2$). Although the fits seem good, the approximated analytic result for $\beta = 2$ is without a doubt not correct.

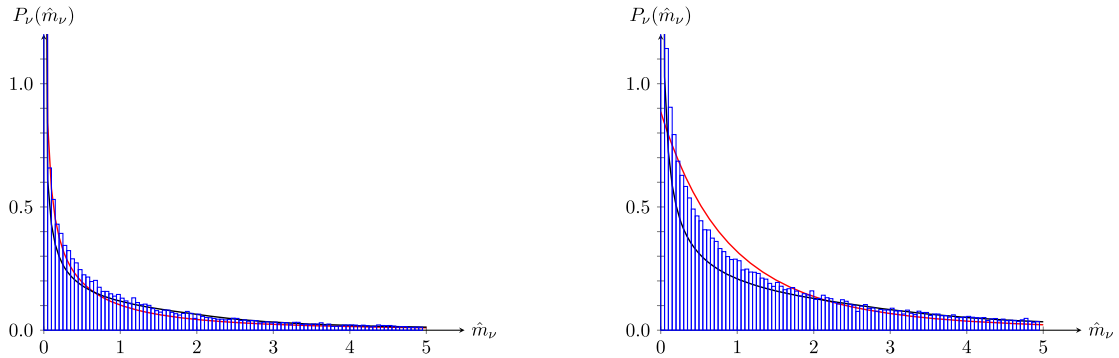


FIG. 3. Comparison between the probability density function at $N = 1$ and the level density at large N for the seesaw ensemble. The red curve corresponds to probability density function, the black curve corresponds to the approximated analytic result while the histogram corresponds to the same numerical results as in Fig. 2. The left panel shows matrices with real elements ($\beta = 1$) while the right panel shows matrices with complex elements ($\beta = 2$).

understanding of the dependence on N . Moreover, this comparison will help show that the analysis of the previous section is most likely wrong even for $\beta = 1$.

In the usual Gaussian ensemble, the comparison is between a Gaussian distribution for the PDF and Wigner's semicircle distribution for the level density at large N . It is clear that the agreement between the two functions is poor. Nevertheless, comparing qualitatively the distributions shows quite strikingly that the support of the normalized level density changes from being noncompact at $N = 1$ to being compact at large N .

For the seesaw ensemble, it is convenient to discuss the real and complex cases separately.

For the complex case ($\beta = 2$), it is already known that the approximated analytic level density at large N is not correct, hence it is only possible to compare the analytic PDF (4.1) with the numerical results in Fig. 2 at large N . The comparison as shown in the right panel of Fig. 3 demonstrates that the PDF at $N = 1$ agrees well with the level density at large N for $\hat{x} \gtrsim 2$ only. This shows qualitatively that the first moment of the level density at large N exists, as expected. Moreover, it is clear that as N increases, the number of normalized singular values around zero changes from a constant for $N = 1$ to possibly become infinite at large N . Finally, as already mentioned, the support stays noncompact even as $N \rightarrow \infty$.

The real case ($\beta = 1$) is more interesting. Since the approximated analytic level density at large N is consistent with the moments and the numerical results, it is possible to compare the analytic results of (4.1) and (4.4) as well as the numerical results of Fig. 2, as shown in the left panel of Fig. 3. A quick glance at Fig. 3 is enough to realize that although the PDF at $N = 1$ and the approximated analytic result for the level density at large N are quite alike, the former is an even better fit to the numerical results for the level density at large N . Hence the normalized level density at large N is well approximated by the PDF at $N = 1$ (4.1)! This observation implies that the normalized level density

does not change much as N increases. It would be interesting to investigate this behavior in more detail.

In summary, the action of first computing the PDF for the seesaw ensemble from the Dirac and Majorana neutrino mass matrices and then obtaining the seesaw ensemble level density at large N is not the same than first obtaining the level densities at large N for the Dirac and Majorana neutrino mass matrices and then computing the seesaw ensemble level density at $N \rightarrow \infty$. The simplifications made by interchanging the order of the steps are too extreme to generate a useful result. From the numerical analysis, one can nevertheless conclude that $\hat{\rho}_{\nu N}^{\beta=1}$ is well approximated by P_ν at $N = 1$ while $\hat{\rho}_{\nu N}^{\beta=2}(\hat{x})$ most likely behaves as $\hat{x}^{-1/2}$ when $\hat{x} \rightarrow 0$ and \hat{x}^{-3} when $\hat{x} \rightarrow \infty$.¹

V. DISCUSSION AND CONCLUSION

This paper investigated the seesaw ensemble originating from the anarchy principle with the help of random matrix theory. The starting assumption was an extended SM with the type I seesaw mechanism, although the same analysis can be done for the type III seesaw mechanism since the light neutrino mass matrix is analogous.² It is shown that the seesaw ensemble joint PDF for the singular values is a complicated function of the singular values and that it decouples from the seesaw ensemble joint PDF for the light neutrino mixing angles and phases which is simply given by the appropriate group (Haar) measure. The fact that the light neutrino mass PDF and the light neutrino mixing matrix PDF are statistically independent is important when comparing with actual neutrino physics since it restricts the

¹The behavior of the PDF around zero suggests that the level density vanishes at zero. If the level density peaks very close to zero before reaching zero as \hat{x} decreases, this feature would be lost in the numerical analysis due to the finite binning.

²Type II seesaw mechanism is trivial since the light neutrino mass matrix would be generated by a Gaussian ensemble.

predictive power of the anarchy principle. Its implications are briefly discussed below.

The asymptotic behavior of the PDF at large mass values suggests that, apart from the average value for complex matrices, none of the positive moments exist. The $N = 1$ case, for which the complicated function simplifies, is studied analytically. The level density at large N is then approximated but it is shown afterwards that the technique used to reach a simplified analytic result is most likely wrong. Moreover, the numerical behavior of the level density at large N found here does not seem to match with the answer found in [5] where an ansatz was used.

Although the physical case of complex 3×3 matrices will be discussed extensively elsewhere [10], the analysis presented here is sufficient to lay out some physical implications of the seesaw ensemble:

1. The seesaw ensemble prefers three massive light neutrinos, a massless neutrino is forbidden.
2. In the seesaw ensemble, there does not exist a correlation between the light neutrino mass eigenstates and the mixing angles and phases, which results in a loss of predictive power when comparing with observations.
3. This lack of correlation implies that the distribution of light neutrino mixing angles and phases is simply the appropriate Haar measure, as pointed out in [3], and hence near-maximal mixings observed experimentally are highly probable although the connection with the mass eigenstates is lost.
4. The absence of a correlation also implies that the spectrum of light neutrino masses can exhibit the normal hierarchy, the inverted hierarchy or something else (contrary to the claim of [5]), as the mostly electronic neutrino could be the heaviest one, in conflict with observations.

The last two observations can be answered quantitatively with a thorough investigation of the seesaw ensemble PDF for $N = 3$ and $\beta = 2$, an analysis that will appear in subsequent work [10].

ACKNOWLEDGMENTS

The authors would like to thank Patrick Desrosiers for enlightening discussions on the mathematics of random matrix theory. This work is supported by NSERC (Natural Sciences and Engineering Research Council of Canada).

APPENDIX A: MEASURES

This appendix rederives important results on the singular value decompositions of the measures encountered in the computation of the seesaw PDF. Note that these measures were already discussed in [3], yet their derivation relied in part on some heuristic arguments that led to conjectures for matrices of arbitrary dimensions. This appendix presents complete proofs making use of the wedge product technique.

1. Stiefel manifold

In constructing the measures from the singular value decomposition theorem, a recurrent quantity of interest is the volume of Stiefel manifolds [12]. The volume of the Stiefel manifold $\mathcal{V}_N^\beta \equiv \mathcal{V}_{N,N}^\beta$ is a well-known result and is simply given by

$$\text{Vol}(\mathcal{V}_N^\beta) = \int_{U \in \mathcal{V}_N^\beta} U^\dagger dU = \frac{2^N \pi^{\beta N(N+1)/4}}{\prod_{1 \leq i \leq N} \Gamma(\beta i/2)}. \quad (\text{A1})$$

This quantity helps define the group measures that appear in the singular value decomposition.

2. Singular value decomposition

With the knowledge of Stiefel volumes, it is now possible to obtain the measures (3.4) and (3.5) using the power of the wedge product approach. Because of its anticommutativity, the wedge product approach allows us to obtain Jacobians simply. However, since each singular value decomposition is different [14], in the following each of the four cases (real symmetric, complex symmetric, real arbitrary and complex arbitrary) are proved independently. Note that overall minus signs are not important in the wedge product approach since only the absolute value of the Jacobian is of interest.

a. Real and complex symmetric matrix

For a real (complex) $N \times N$ symmetric (denoted by the subscript S) matrix M , the eigen (singular) value decomposition theorem states that there exists a real (complex) matrix U satisfying $U^\dagger U = \mathbb{1}$ such that $M = UDU^T$ where D is real diagonal, i.e. $D = \text{diag}(\lambda_1, \dots, \lambda_N)$, with $\lambda_N > \lambda_{N-1} > \dots > \lambda_1$ and $\lambda_1 > 0$ ($\lambda_1 > -\infty$). Thus, one has

$$dM = d(UDU^T) = dUDU^T + UdDU^T + UDdU^T. \quad (\text{A2})$$

Multiplying (A2) by U^\dagger on the left and by U^* on the right leads to

$$\begin{aligned} U^\dagger dMU^* &= (U^\dagger dU)D + dD + D(U^\dagger dU)^T \\ &= (U^\dagger dU)D + dD - D(U^\dagger dU)^* \\ &= \begin{cases} d\lambda_i + \lambda_i[(U^\dagger dU)_{ii} - (U^\dagger dU)_{ii}^*] & \text{for } i = j \\ \lambda_j(U^\dagger dU)_{ij} - \lambda_i(U^\dagger dU)_{ij}^* & \text{for } i \neq j \end{cases} \end{aligned} \quad (\text{A3})$$

where the second and last equalities come from $U^\dagger dU = d(U^\dagger U) - dU^\dagger U = -(U^\dagger dU)^\dagger$. From (3.9), the wedge product on the lhs of (A3) is simply the wedge product of dM , i.e. of its independent elements, hence using the wedge product on the rhs for a real symmetric matrix one gets

$$\begin{aligned}
 (dM) &\equiv \bigwedge_{1 \leq i < j \leq N} dM_{ij} \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i < j \leq N} (\lambda_j - \lambda_i) (U^\dagger dU)_{ij} \\
 &= \prod_{1 \leq i < j \leq N} (\lambda_j - \lambda_i) \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i < j \leq N} (U^\dagger dU)_{ij},
 \end{aligned}$$

while for a complex symmetric matrix one obtains

$$\begin{aligned}
 (dM) &\equiv \bigwedge_{1 \leq i < j \leq N} d\text{Re}M_{ij} \wedge d\text{Im}M_{ij} \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \wedge 2\lambda_i \text{Im}(U^\dagger dU)_{ii} \bigwedge_{1 \leq i < j \leq N} (\lambda_j - \lambda_i) \\
 &\quad \times \text{Re}(U^\dagger dU)_{ij} \wedge (\lambda_j + \lambda_i) \text{Im}(U^\dagger dU)_{ij} \\
 &= 2^N \prod_{1 \leq i \leq N} \lambda_i \prod_{1 \leq i < j \leq N} (\lambda_j^2 - \lambda_i^2) \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i < j \leq N} (U^\dagger dU)_{ij}.
 \end{aligned}$$

Hence, the measure is

$$dM = c_{SN}^\beta \prod_{1 \leq i < j \leq N} |\lambda_i^\beta - \lambda_j^\beta| \prod_{1 \leq i \leq N} \lambda_i^{\beta-1} d\lambda_i U^\dagger dU, \quad (\text{A4})$$

where the normalization constant is

$$c_{SN}^\beta = \frac{2^{(\beta-1)N}}{2^N N!} = \frac{1}{2^{(2-\beta)N} N!}, \quad (\text{A5})$$

and the integration region is extended to $-\infty < \lambda_i < \infty$ ($0 < \lambda_i < \infty$) for all eigen (singular) values [hence the factor of $N!$ in the denominator of (A5)]. The extra factor of 2^N in the denominator of (A5) accounts for the remaining freedom in the eigen (singular) value decomposition $M = UDU^T$ where U is replaced by US with $S = \text{diag}(\pm 1, \dots, \pm 1)$. This factor implies the integration region for the group measure is the full region with the Stiefel volume mentioned above.

b. Real and complex arbitrary matrix

The proof of the measure for an arbitrary (denoted by the subscript A) matrix is mostly equivalent. For a real

$$\begin{aligned}
 (dM) &\equiv \bigwedge_{1 \leq i, j \leq N} d\text{Re}M_{ij} \wedge d\text{Im}M_{ij} \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \wedge \lambda_i \text{Im}[(U^\dagger dU)_{ii} - (V^\dagger dV)_{ii}] \bigwedge_{1 \leq i \neq j \leq N} [\lambda_j \text{Re}(U^\dagger dU)_{ij} - \lambda_i \text{Re}(V^\dagger dV)_{ij}] \wedge [\lambda_j \text{Im}(U^\dagger dU)_{ij} - \lambda_i \text{Im}(V^\dagger dV)_{ij}] \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \wedge \lambda_i \text{Im}[(U^\dagger dU)_{ii} - (V^\dagger dV)_{ii}] \bigwedge_{1 \leq i < j \leq N} [\lambda_j \text{Re}(U^\dagger dU)_{ij} - \lambda_i \text{Re}(V^\dagger dV)_{ij}] \wedge [\lambda_j \text{Re}(V^\dagger dV)_{ij} - \lambda_i \text{Re}(U^\dagger dU)_{ij}] \\
 &\quad \bigwedge_{1 \leq i < j \leq N} [\lambda_j \text{Im}(U^\dagger dU)_{ij} - \lambda_i \text{Im}(V^\dagger dV)_{ij}] \wedge [\lambda_i \text{Im}(U^\dagger dU)_{ij} - \lambda_j \text{Im}(V^\dagger dV)_{ij}] \\
 &= \prod_{1 \leq i \leq N} \lambda_i \prod_{1 \leq i < j \leq N} (\lambda_j^2 - \lambda_i^2)^2 \bigwedge_{1 \leq i \leq N} d\lambda_i \wedge \text{Im}[(U^\dagger dU)_{ii} - (V^\dagger dV)_{ii}] \bigwedge_{1 \leq i < j \leq N} (U^\dagger dU)_{ij} \wedge (V^\dagger dV)_{ij}.
 \end{aligned}$$

(complex) $N \times N$ arbitrary matrix M , the singular value decomposition theorem implies that there exist real (complex) matrices U and V satisfying $U^\dagger U = 1$ and $V^\dagger V = 1$ such that $M = UDV^\dagger$ where D is real diagonal, i.e. $D = \text{diag}(\lambda_1, \dots, \lambda_N)$, with $\lambda_N > \lambda_{N-1} > \dots > \lambda_1 > 0$. Therefore, one has

$$dM = d(UDV^\dagger) = dUDV^\dagger + UdDV^\dagger + UdV^\dagger dV. \quad (\text{A6})$$

Multiplying (A2) by U^\dagger on the left and by V on the right leads to

$$\begin{aligned}
 U^\dagger dMV &= (U^\dagger dU)D + dD + D(V^\dagger dV)^\dagger \\
 &= (U^\dagger dU)D + dD - D(V^\dagger dV) \\
 &= \begin{cases} d\lambda_i + \lambda_i[(U^\dagger dU)_{ii} - (V^\dagger dV)_{ii}] & \text{for } i = j \\ \lambda_j(U^\dagger dU)_{ij} - \lambda_i(V^\dagger dV)_{ij} & \text{for } i \neq j \end{cases},
 \end{aligned} \quad (\text{A7})$$

where the last two equalities come from $V^\dagger dV = d(V^\dagger V) - dV^\dagger V = -(V^\dagger dV)^\dagger$. Again, the wedge product on the lhs of (A7) is simply the wedge product of dM . Hence using the wedge product on the rhs for a real arbitrary matrix one obtains

$$\begin{aligned}
 (dM) &\equiv \bigwedge_{1 \leq i, j \leq N} dM_{ij} \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i \neq j \leq N} [\lambda_j(U^\dagger dU)_{ij} - \lambda_i(V^\dagger dV)_{ij}] \\
 &= \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i < j \leq N} [\lambda_j(U^\dagger dU)_{ij} - \lambda_i(V^\dagger dV)_{ij}] \\
 &\quad \wedge [\lambda_j(V^\dagger dV)_{ij} - \lambda_i(U^\dagger dU)_{ij}] \\
 &= \prod_{1 \leq i < j \leq N} (\lambda_j^2 - \lambda_i^2) \bigwedge_{1 \leq i \leq N} d\lambda_i \bigwedge_{1 \leq i < j \leq N} (U^\dagger dU)_{ij} \\
 &\quad \wedge (V^\dagger dV)_{ij},
 \end{aligned}$$

while for a complex arbitrary matrix one gets

Therefore, the final measure becomes

$$dM = c_{AN}^\beta \prod_{1 \leq i < j \leq N} |\lambda_i^2 - \lambda_j^2|^\beta \prod_{1 \leq i \leq N} \lambda_i^{\beta-1} d\lambda_i (U^\dagger dU)' (V^\dagger dV)', \quad (\text{A8})$$

where the normalization constant is

$$c_{AN}^\beta = \frac{1}{2^{(2-\beta)N} N!}, \quad (\text{A9})$$

and the integration region is over all non-negative singular values, i.e. $0 < \lambda_i < \infty$. Again, this accounts for the factor of $N!$ in the denominator of (A9). The other factor of $2^{(2-\beta)N}$ in the denominator of (A9) accounts for the remaining freedom in the singular value decomposition $M = UDV^\dagger$ where U and V are replaced respectively by US and VS with $S = \text{diag}(\pm 1, \dots, \pm 1)$. For a real arbitrary matrix, this factor implies the group measure is integrated over the full region with the Stiefel volume mentioned above. For a complex arbitrary matrix, the remaining freedom is already taken care of by the fact that not all diagonal elements of $U^\dagger dU$ and $V^\dagger dV$ appear in the measure (A8) but only the specific combination $\text{Im}[(U^\dagger dU)_{ii} - (V^\dagger dV)_{ii}]$ does. This observation is denoted by a prime on $U^\dagger dU$ and $V^\dagger dV$. Although it is counter-intuitive to have more integration parameters on the rhs than the lhs of (A8), it is always possible to introduce the missing integration variables [for example, $\text{Im}(U^\dagger dU)_{ii}$] and compensate by dividing the measure by the appropriate volume, i.e. $(2\pi)^{(\beta-1)N}$. This trick allows integrating over the full Stiefel volumes of both U and V in the complex case also, as in Sec. III A.

APPENDIX B: CIRCULAR LAW

In this appendix the different circular laws for Gaussian ensembles are rederived from the Coulomb gas approach.

1. Level density

The level density $\rho(x)$ for real ($\beta = 1$) and complex ($\beta = 2$), symmetric (S) and arbitrary (A), matrices M with usual Gaussian ensembles for the singular values λ_i given by

$$P_S(\lambda) d\lambda \propto \prod_{1 \leq i < j \leq N} |\lambda_i^\beta - \lambda_j^\beta| \prod_{1 \leq i \leq N} |\lambda_i|^{\beta-1} e^{-\lambda_i^\beta} d\lambda_i,$$

$$P_A(\lambda) d\lambda \propto \prod_{1 \leq i < j \leq N} |\lambda_i^2 - \lambda_j^2|^\beta \prod_{1 \leq i \leq N} \lambda_i^{\beta-1} e^{-\lambda_i^2} d\lambda_i,$$

[compare to (3.6)] can be computed from the associated Coulomb-like gas with Hamiltonians

$$H_S = -\frac{1}{2} \int dx dy \rho_{SN}^\beta(x) \rho_{SN}^\beta(y) \ln |x^\beta - y^\beta| - (\beta - 1) \times \int dx \rho_{SN}^\beta(x) \ln |x| + \int dx \rho_{SN}^\beta(x) x^2,$$

$$H_A = -\frac{1}{2} \int dx dy \rho_{AN}^\beta(x) \rho_{AN}^\beta(y) \ln |x^2 - y^2| - \left(1 - \frac{1}{\beta}\right) \int dx \rho_{AN}^\beta(x) \ln(x) + \frac{1}{\beta} \int dx \rho_{AN}^\beta(x) x^2. \quad (\text{B1})$$

Extremizing (B1) with respect to the level density, subject to the normalization constraint $\int dx \rho(x) = N$, leads to

$$\xi_S = - \int dy \rho_{SN}^\beta(y) \ln |x^\beta - y^\beta| - (\beta - 1) \ln |x| + x^2,$$

$$\xi_A = - \int dy \rho_{AN}^\beta(y) \ln |x^2 - y^2| - \left(1 - \frac{1}{\beta}\right) \ln(x) + \frac{1}{\beta} x^2, \quad (\text{B2})$$

where ξ is the constant Lagrange multiplier. Deriving (B2) with respect to x gives

$$0 = -\beta x^{\beta-1} \mathcal{P} \int dy \frac{\rho_{SN}^\beta(y)}{x^\beta - y^\beta} - \frac{\beta-1}{x} + 2x,$$

$$0 = -2x \mathcal{P} \int dy \frac{\rho_{AN}^\beta(y)}{x^2 - y^2} - \left(1 - \frac{1}{\beta}\right) \frac{1}{x} + \frac{2}{\beta} x, \quad (\text{B3})$$

where \mathcal{P} denotes the principal value. At this point, it is convenient to rescale the variables such that the thermodynamic limit with N large can be easily taken. This is done with the help of $x = \sqrt{N} \hat{x}$ and $\rho(x) = \sqrt{N} \hat{\rho}(\hat{x})$ which change (B3) to

$$0 = -\beta \hat{x}^{\beta-1} \mathcal{P} \int d\hat{y} \frac{\hat{\rho}_{SN}^\beta(\hat{y})}{\hat{x}^\beta - \hat{y}^\beta} - \frac{\beta-1}{N \hat{x}} + 2\hat{x},$$

$$0 = -2\hat{x} \mathcal{P} \int d\hat{y} \frac{\hat{\rho}_{AN}^\beta(\hat{y})}{\hat{x}^2 - \hat{y}^2} - \left(1 - \frac{1}{\beta}\right) \frac{1}{N \hat{x}} + \frac{2}{\beta} \hat{x}, \quad (\text{B4})$$

and the normalization condition to $\int d\hat{x} \hat{\rho}(\hat{x}) = 1$. In the thermodynamic limit, the second term in each equation of (B4) is negligible and can be discarded. From now on the normalized level density is easily obtained from the usual treatment and is given by

$$\hat{\rho}_{SN}^\beta(\hat{x}) = \frac{2}{\pi} \sqrt{\beta - \hat{x}^2} \quad (\beta - 2) \sqrt{\beta} \leq \hat{x} \leq \sqrt{\beta},$$

$$\hat{\rho}_{AN}^\beta(\hat{x}) = \frac{2}{\beta \pi} \sqrt{2\beta - \hat{x}^2} \quad 0 \leq \hat{x} \leq \sqrt{2\beta}. \quad (\text{B5})$$

Note that the small differences between the circular laws obtained here and the usual ones come from the different normalizations used.

- [1] K. A. Olive *et al.* (Particle Data Group Collaboration), Review of Particle Physics, *Chin. Phys. C* **38**, 090001 (2014); M. C. Gonzalez-Garcia, M. Maltoni, and T. Schwetz, Updated fit to three neutrino mixing: Status of leptonic CP violation, *J. High Energy Phys.* **11** (2014) 052.
- [2] L. J. Hall, H. Murayama, and N. Weiner, Neutrino Mass Anarchy, *Phys. Rev. Lett.* **84**, 2572 (2000).
- [3] N. Haba and H. Murayama, Anarchy and hierarchy, *Phys. Rev. D* **63**, 053010 (2001).
- [4] A. de Gouvea and H. Murayama, Statistical test of anarchy, *Phys. Lett. B* **573**, 94 (2003); Neutrino mixing anarchy: Alive and kicking, *Phys. Lett. B* **747**, 479 (2015); J. Heeck, Seesaw parametrization for n right-handed neutrinos, *Phys. Rev. D* **86**, 093023 (2012).
- [5] Y. Bai and G. Torroba, Large N ($= 3$) neutrinos and random matrix theory, *J. High Energy Phys.* **12** (2012) 026.
- [6] X. Lu and H. Murayama, Neutrino mass anarchy and the universe, *J. High Energy Phys.* **08** (2014) 101.
- [7] M. Mehta, *Random Matrices* (Elsevier Science, Amsterdam, 2004).
- [8] R. Muirhead, *Aspects of Multivariate Statistical Theory* (Wiley, New York, 2009).
- [9] P. Forrester, *Log-Gases and Random Matrices (LMS-34)* (Princeton University Press, Princeton, 2010).
- [10] J.-F. Fortin, N. Giasson, and L. Marleau (to be published).
- [11] P. Minkowski, $\mu \rightarrow e\gamma$ at a rate of one out of 10^9 muon decays?, *Phys. Lett.* **67B**, 421 (1977).
- [12] T. Ratnarajah, R. Vaillancourt, and M. Alvo, Jacobians and hypergeometric functions in complex multivariate analysis, *Canad. Appl. Math. Quart.* **12**, 213 (2004).
- [13] C. Spengler, M. Huber, and B. C. Hiesmayr, A composite parameterization of unitary groups, density matrices and subspaces, *J. Phys. A* **43**, 385306 (2010), Composite parameterization and Haar measure for all unitary and special unitary groups, *J. Math. Phys.* **53**, 013501 (2012).
- [14] T. Ratnarajah and R. Vaillancourt, Complex singular wishart matrices and applications, *Comput. Math. Appl.* **50**, 399 (2005).