Chin. Ann. Math.
38B(2), 2017, 539–562
DOI: 10.1007/s11401-017-1082-9

Chinese Annals of Mathematics, Series B

Variational Analysis of Toda Systems*

Andrea MALCHIODI[1]

(*To Haïm with admiration and gratitude*)

Abstract The author surveys some recent progress on the Toda system on a two-dimensional surface Σ, arising in models from self-dual non-abelian Chern-Simons vortices, as well as in differential geometry. In particular, its variational structure is analysed, and the role of the topological join of the barycentric sets of Σ is shown.

Keywords Geometric PDEs, Variational methods, Min-max schemes
2000 MR Subject Classification 35B33, 35J35, 53A30, 53C21

1 Introduction

The Toda system consists of coupled Liouville equations of the type

$$-\frac{1}{2}\Delta u_i(x) = \sum_{j=1}^{N} a_{ij}\mathrm{e}^{u_j(x)}, \quad x \in \Sigma,\ i = 1, \cdots, N, \tag{1.1}$$

where $\Delta = \Delta_g$ is Laplace-Beltrami operator and $A = (a_{ij})_{ij}$ is the Cartan matrix of $SU(N+1)$,

$$A = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

This system arises in the study of self-dual non-abelian Chern-Simons models (see [22, 44, 45]) (also for further details and an up-to-date set of references), and the right-hand side of the equations might contain singular sources corresponding to vortices, namely points where the wave function appearing in the physical model vanishes. The system also has an interest in geometry, as it describes to the Frenet frame of (possibly ramificated) holomorphic curves in $\mathbb{CP}^n$ (see [8, 10, 16, 26]).

The following non-homogeneous version with two components was extensively studied on compact, boundary-less Riemannian surfaces (Σ, g):

$$\begin{cases} -\Delta u_1 = 2\rho_1\Big(\dfrac{h_1\mathrm{e}^{u_1}}{\int_\Sigma h_1\mathrm{e}^{u_1}\mathrm{d}V_g} - 1\Big) - \rho_2\Big(\dfrac{h_2\mathrm{e}^{u_2}}{\int_\Sigma h_2\mathrm{e}^{u_2}\mathrm{d}V_g} - 1\Big), \\ -\Delta u_2 = 2\rho_2\Big(\dfrac{h_2\mathrm{e}^{u_2}}{\int_\Sigma h_2\mathrm{e}^{u_2}\mathrm{d}V_g} - 1\Big) - \rho_1\Big(\dfrac{h_1\mathrm{e}^{u_1}}{\int_\Sigma h_1\mathrm{e}^{u_1}\mathrm{d}V_g} - 1\Big). \end{cases} \tag{1.2}$$

Manuscript received June 11, 2015. Revised October 13, 2015.
[1]Scuola Normale Superiore, Piazza dei Cavalieri 7, 50126 Pisa, Italy. E-mail: andrea.malchiodi@sns.it
*The author is supported by the project PRIN 2015 2015KB9WPT_001 and is a member of INdAM.

Here h_1, h_2 are smooth positive functions on Σ and ρ_1, ρ_2 are real parameters. Flat tori might model for example periodic physical systems in the plane.

(1.2) has variational structure, and the corresponding Euler functional $J_\rho : H^1(\Sigma) \times H^1(\Sigma) \to \mathbb{R}$ has the expression

$$J_\rho(u_1, u_2) = \int_\Sigma Q(u_1, u_2)\, \mathrm{d}V_g + \sum_{i=1}^{2} \rho_i \Big(\int_\Sigma u_i \mathrm{d}V_g - \log \int_\Sigma h_i \mathrm{e}^{u_i} \mathrm{d}V_g \Big), \quad \rho = (\rho_1, \rho_2), \tag{1.3}$$

where $Q(u_1, u_2)$ is the positive-definite quadratic form as follows:

$$Q(u_1, u_2) = \frac{1}{3}(|\nabla u_1|^2 + |\nabla u_2|^2 + \nabla u_1 \cdot \nabla u_2). \tag{1.4}$$

It is well-known that $H^1(\Sigma)$ embeds into any L^p space, and that indeed the embedding can be pushed to exponential class via the Moser-Trudinger inequality. Concerning the functional J_ρ the sharp inequality for the Toda system was found in [26].

Theorem 1.1 (see [26]) *For $\rho = (\rho_1, \rho_2)$ the functional $J_\rho : H^1(\Sigma) \times H^1(\Sigma)$ is bounded from below if and only if both ρ_1 and ρ_2 satisfy $\rho_i \le 4\pi$.*

By the latter theorem we have that when both $\rho_1, \rho_2 < 4\pi$ the functional J_ρ is coercive, and solutions can be found by minimization via the direct methods of the calculus of variations (see also [28] for the case when $\max_i \rho_i = 4\pi$, when the energy is bounded below but non-coercive). When one of the ρ_i's exceeds 4π the energy becomes unbounded from below, and solutions have to be found as saddle points. One result in this direction is as follows.

Theorem 1.2 (see [36]) *Suppose that m is a positive integer, and let $h_1, h_2 : \Sigma \to \mathbb{R}$ be smooth positive functions. Then for $\rho_1 \in (4\pi m, 4\pi(m+1))$ and for $\rho_2 < 4\pi$, (1.2) is solvable.*

Remark 1.1 The case $m = 1$ in Theorem 1.2 was proved in [25] for surfaces with positive genus. The assumption $\rho_1 \notin 4\pi\mathbb{N}$ is due to compactness reasons (see Section 2).

To describe the general ideas beyond the proof of Theorem 1.2, we recall first the mean field or Liouville equation:

$$-\Delta u = 2\widetilde{\rho}\Big(\frac{h(x)\mathrm{e}^u}{\int_\Sigma h(x)\mathrm{e}^u dV_g} - 1 \Big), \tag{1.5}$$

where $\widetilde{\rho} \in \mathbb{R}$ and $h : \Sigma \to \mathbb{R}$ is smooth and positive. The interest in (1.5) arises from the abelian version of (1.2) and, in geometry, from the problem of conformally prescribing the Gaussian curvature of a compact surface (see [1]).

One approach to attack the existence problem for (1.5) relies on computing the Leray-Schauder degree of the equation (see [14, 29]) (more comments on this approach will be given in the next section). Another one exploits the variational structure of the problem, via the Euler functional $I_{\widetilde{\rho}} : H^1(\Sigma) \to \mathbb{R}$,

$$I_{\widetilde{\rho}}(u) = \frac{1}{2} \int_\Sigma |\nabla_g u|^2 \mathrm{d}V_g + 2\widetilde{\rho}\Big(\int_\Sigma u \mathrm{d}V_g - \log \int_\Sigma h(x)\mathrm{e}^u \mathrm{d}V_g \Big), \quad u \in H^1(\Sigma). \tag{1.6}$$

The counterpart of Theorem 1.1 in the scalar case is the classical Moser-Trudinger inequality ($\overline{u}$ stands for the average of u on Σ)

$$\log \int_\Sigma \mathrm{e}^{(u - \overline{u})} \mathrm{d}V_g \le C + \frac{1}{16\pi} \int_\Sigma |\nabla u|^2 \mathrm{d}V_g, \tag{1.7}$$

which gives coercivity of $I_{\widetilde{\rho}}$ for $\rho < 4\pi$ (see [18, 40] for the borderline case). The supercritical case $\rho > 4\pi$ can be treated through improvements of the latter inequality under suitable conditions on the function u (see [13]). Roughly, the improvement states that if the function e^u spreads into two separate regions of Σ, then the constant in (1.7) can be nearly halved. In [19], (1.5) was studied for $\widetilde{\rho} \in (4\pi, 8\pi)$, and the improvement in [13] was used to show that if $I_{\widetilde{\rho}}$ is large negative, then the probability measure $\frac{e^u}{\int_\Sigma e^u dV_g}$ has to concentrate near a single point of Σ. This fact was used jointly with a variational scheme to prove existence of critical points of saddle point type when Σ has positive genus (see [43] for a different argument on the flat torus, using the mountain-pass theorem).

This strategy was then pursued in [21] (for the prescribed Q-curvature problem in four dimension) and in [20] to treat the case $\widetilde{\rho} \in (4k\pi, 4(k+1)\pi)$. An extension of the argument in [13, 19], with a more involved topological construction, allowed to show that for low energy the measure $\frac{e^u}{\int_\Sigma e^u dV_g}$ concentrates near at most k points of the surface. This induces to consider the family Σ_k of formal sums

$$\Sigma_k = \Big\{ \sum_{i=1}^{k} t_i \delta_{x_i} : \sum_{i=1}^{k} t_i = 1,\ t_i \geq 0,\ x_i \in \Sigma,\ \forall\, i = 1, \cdots, k \Big\}, \tag{1.8}$$

called the barycentric sets of Σ of order k. The above set of measures, which is naturally endowed with the weak topology of distributions, does not have a smooth structure for $k \geq 2$ (while Σ_1 is homeomorphic to Σ), that is, it is a stratified set, namely union of open manifolds of different dimensions (see [27] for further characterizations, especially of topological type). The basic property used in [20–21] is that Σ_k is non-contractible, which allows to define proper min-max schemes to attack the existence problem. In [35], it was also used to deduce the degree-counting formula from [14] with a different approach. We also mention the role of these sets for the study of the Yamabe equation in Euclidean domains (see [2]).

In [36], it was shown that this latter approach can be extended to study (1.2) in the situation described in Theorem 1.2. When $\rho_1 \in (4k\pi, 4(k+1)\pi)$, $k \in \mathbb{N}$, and $\rho_2 < 4\pi$ the Euler-Lagrange energy J_ρ is virtually (even though not literally) coercive in the second component u_2, and the set Σ_k again appears in the distributional description of the function e^{u_1} when J_ρ is low enough. Using the non-contractibility of Σ_k, one can then prove Theorem 1.2 using variational techniques.

We are interested here in the situation when both the ρ_i's exceed the threshold coercivity value 4π. Using improved inequalities in the spirit of [13], it is possible to prove that if $\rho_1 < 4(k+1)\pi$, $\rho_2 < 4(l+1)\pi$, $k, l \in \mathbb{N}$, and if $J_\rho(u_1, u_2)$ is sufficiently low, then either e^{u_1} is close to Σ_k or e^{u_2} is close to Σ_l in the distributional sense. This (non-mutually exclusive) alternative can be expressed in term of the topological join of Σ_k and Σ_l. Recall that, given two topological spaces A and B, their join $A * B$ is defined as the family of elements of the form (see [23])

$$A * B = \frac{\{(a, b, s) :\ a \in A,\ b \in B,\ s \in [0, 1]\}}{E}, \tag{1.9}$$

where E is an equivalence relation given by

$$(a_1, b, 1) \stackrel{E}{\sim} (a_2, b, 1), \quad \forall\, a_1, a_2 \in A,\ b \in B \quad \text{and} \quad (a, b_1, 0) \stackrel{E}{\sim} (a, b_2, 0), \quad \forall\, a \in A,\ b_1, b_2 \in B.$$

This construction allows to map low sublevels of J_ρ into $\Sigma_k * \Sigma_l$, with the join parameter s expressing whether distributionally e^{u_1} is closer to Σ_k or whether e^{u_2} is closer to Σ_l.

However, as for the scalar problem (1.5) the above description is somehow optimal, it is no more the case for (1.2). The optimality of the description of low-energy levels of $I_{\widetilde{\rho}}$ by means of the Σ_k's has to be understood in the following sense. When $\widetilde{\rho} \in (4k\pi, 4(k+1)\pi)$, for any measure $\sigma \in \Sigma_k$ there exists a test function $\varphi_{\lambda,\sigma}$, depending on a large parameter $\lambda > 0$ with the following properties:

$$e^{\varphi_{\lambda,\sigma}} \rightharpoonup \sigma, \quad I_{\widetilde{\rho}}(\varphi_{\lambda,\sigma}) \to -\infty \quad \text{as } \lambda \to +\infty,$$

with the second property holding uniformly in the choice of $\sigma \in \Sigma_k$. This means that copies of Σ_k can naturally be embedded into arbitrary low sublevels of $I_{\widetilde{\rho}}$. In this way, using also the previous statements, one finds a way to go back and forth from large-negative sublevels of $I_{\widetilde{\rho}}$ to the set of measures Σ_k.

Turning to (1.2), things get more complicated, due to the interaction form $Q(u_1, u_2)$. Looking at its structure, it clearly appears that when the gradients of u_1 and u_2 point in the same direction Q gets larger. This means that if u_1 and u_2 are peaked near the same point a higher energy is expected, which puts extra constraints in the choice of the test functions. In some cases this problem can be overcome: By restricting the location of the peaks of the two components u_1, u_2 to disjoint curves $\gamma_1, \gamma_2 \subseteq \Sigma$, the following result was obtained in [7].

Theorem 1.3 (see [7]) *Suppose that $\rho_i \notin 4\pi\mathbb{N}$ for both $i = 1, 2$ and that Σ has positive genus. Then* (1.2) *is solvable.*

The assumption on the genus of Σ is used to construct global maps from Σ into γ_i, and then their push-forwards from Σ_k (resp. Σ_l) into $(\gamma_1)_k$ (resp. $(\gamma_2)_l$) (the barycentric sets of the γ_i's), and then reason only in terms of the topological join of $(\gamma_1)_k$ and $(\gamma_2)_l$. Some counterpart of this argument for the singular Liouville equation can be found in [3].

In general, one would need to understand in more depth the interaction between the two components when they concentrate near the same point. We will see that one has to take into account not only of the location of the concentration points, but also of the scale of concentration. One can indeed show for example that when both ρ_i's belong to $(4\pi, 8\pi)$, $J_\rho(u_1, u_2)$ low enough implies that either e^{u_1}, e^{u_2} are concentrated at different points, or that if they are concentrated at the same point but with different scales of concentration (see [37]).

This argument relies on improved Moser-Trudinger inequalities which are different in spirit from those in [13], and which have the feature of being scaling-invariant. They can be applied to functions that are arbitrarily concentrated and that, at a macroscopic level, look just like Dirac masses. A different improved inequality, but with the same scaling-invariant feature, was proved in [24]. We refer to Section 3 for details, just remarking here that it applies to cases when one component is much more concentrated than the other. This has some relation to improved inequalities for the Singular Liouville equation: A version of (1.5) with singular sources (in the form of Dirac masses) on the right-hand side, representing either conical points on surfaces, or vortex points from the physical point of view. The relation to this latter problem can be seen from the right-hand side of (1.2): A highly concentrated function appears as a Dirac delta when looking at the scale of the less concentrated component. These improved inequalities have the effect of removing suitable subsets in the topological join $\Sigma_k * \Sigma_l$ (see Sections 3–4 for details). One can then prove the following result via min-max theory.

Theorem 1.4 (see [24, 37]) *Let h_1, h_2 be two positive smooth functions, and let Σ be any* *compact surface. Suppose that* $\rho_1 \in (4k\pi, 4(k+1)\pi),\ k \in \mathbb{N}$ *and* $\rho_2 \in (4\pi, 8\pi)$. *Then* (1.2) *has a solution.*

The existence problem for the case of general parameters and genus is still open. We hope that the topological join construction might still play a role. An interesting variant of (1.2) regards the presence of singular sources on the right-hand sides of the equations. For this case the progress is still limited (see [5, 7, 11] in the scalar case) for some particular situations.

The plan of the paper is the following. In Section 2, some compactness results are presented, showing the role of the multiples of 4π: Applications to min-max constructions are also discussed. In Section 3, some improved inequalities are shown, both at a macroscopic and at a scaling-invariant level. Section 4 treats the weighted barycentric sets and the topological join construction. Finally, in Section 5, some test functions are constructed, which allow to define suitable min-max schemes to prove existence of solutions.

Some Notations Given points $x, y \in \Sigma$, $d(x, y)$ will stand for the metric distance between x and y on Σ. Similarly, for any $p \in \Sigma$, $\Omega, \Omega' \subseteq \Sigma$, we set

$$d(p, \Omega) = \inf\{d(p, x) : x \in \Omega\}, \quad d(\Omega, \Omega') = \inf\{d(x, y) : x \in \Omega,\ y \in \Omega'\}.$$

The symbol $B_s(p)$ stands for the open metric ball of radius s and center p, and the complement of a set Ω in Σ will be denoted by Ω^c.

Given a function $u \in L^1(\Sigma)$ and $\Omega \subset \Sigma$, the average of u on Ω is denoted by the symbol

$$\fint_\Omega u \, \mathrm{d}V_g = \frac{1}{|\Omega|} \int_\Omega u \, \mathrm{d}V_g.$$

We denote by $\overline{u}$ the average of u in Σ. Since we are assuming $|\Sigma| = 1$, we have

$$\overline{u} = \int_\Sigma u \, \mathrm{d}V_g = \fint_\Sigma u \, \mathrm{d}V_g.$$

The sub-levels of the functional J_ρ will be indicated as

$$J_\rho^a := \{u = (u_1, u_2) \in H^1(\Sigma) \times H^1(\Sigma) : J(u_1, u_2) \le a\}.$$

Throughout this paper, the letter C will stand for large constants which are allowed to vary among different formulas or even within the same lines. We denote $\mathcal{M}(\Sigma)$ the set of all Radon measures on Σ, and introduce a norm by using duality versus Lipschitz functions, that is, we set

$$\|\mu\|_{\mathrm{Lip}'(\Sigma)} = \sup_{\|f\|_{\mathrm{Lip}(\Sigma)} \le 1} \left| \int_\Sigma f \, \mathrm{d}\mu \right|, \quad \mu, \nu \in \mathcal{M}(\Sigma). \tag{1.10}$$

We denote by $\mathbf{d}$ the corresponding distance, which is known as the Kantorovich-Rubinstein distance.

2 Analytic Aspects of the Problem

In this section, we collect some useful compactness results that mainly arise from a concentration-compactness alternative. We discuss first the scalar case, then the vector case, and finally we turn to applications to min-max theory.

2.1 Compactness properties of (1.5)

In this subsection, we describe the analytic aspects of the problem, especially for what concerns the compactness properties. The first result in this direction was proved in [9], concerning the scalar Liouville equation in a bounded domain Ω of the Euclidean plane. For a smooth positive function $h(x)$ and a real-valued sequence of positive numbers ρ_n we consider the following problem:

$$\begin{cases} -\Delta u_n = 2\rho_n h(x) \mathrm{e}^{u_n} \quad \text{in } \Omega, \\ \displaystyle\int_\Omega \mathrm{e}^{u_n} \mathrm{d}x \leq C, \quad |\log \rho_n| \leq C. \end{cases} \tag{2.1}$$

Define then the blow-up set S as

$$S = \{x \in \Omega : \exists x_n \to x \text{ such that } u_n(x_n) \to +\infty\}.$$

Theorem 2.1 (see [9]) *Suppose that $\Omega \subseteq \mathbb{R}^2$ is a bounded domain, and consider a sequence of solutions of* (2.1)*. Then, up to a subsequence, one of the following three possibilities holds true:*

(i) *u_n is bounded in $L^\infty_{\mathrm{loc}}(\Omega)$.*

(ii) *$u_n \to -\infty$ on every compact set of Ω.*

(iii) *The blow-up set S of (u_n) is finite, $u_n \to -\infty$ on the compact sets of $\Omega \setminus S$ and moreover*

$$\rho_n h(x) \mathrm{e}^{u_n} \rightharpoonup \sum_{x_i \in S} \beta_i \delta_{x_i}$$

with $\beta_i \geq 2\pi$ for every $x_i \in S$.

This result was proved using potential theory jointly with Jensen's inequality. Later, it was specialized by Li and Shafrir in the following sense.

Theorem 2.2 (see [30]) *If the case* (iii) *occurs in Theorem* 2.1*, then β_i is a positive multiple of 4π for every $x_i \in S$.*

One of the main extra ingredients in the proof of the latter result is the blow-up analysis of solutions. The u_n's are rescaled near its local maxima. Using the dilation invariance of the limit equation

$$-\Delta u = 2\overline{\rho}_1 \mathrm{e}^u \quad \text{in } \mathbb{R}^2, \tag{2.2}$$

it is proved that in the limit a scaled u_n solves (2.2) and has a uniformly bounded global maximum at zero. Solutions of (2.2) with such property were classified in [13]. It was proved that, up to a dilation, the limit profile has the form

$$U_0(x) = 2 \log \frac{1}{1 + |x|^2} + C_{\rho_0} \tag{2.3}$$

for some explicit constant $C_{\rho_0} \in \mathbb{R}$. Such functions have the property that $\rho_0 \int_{\mathbb{R}^2} \mathrm{e}^{U_0} \mathrm{d}x = 4\pi$ independently of ρ_0, which justifies the appearance of the multiples of 4π in Theorem 2.2. Some delicate extra work (relying on Harnack-type inequalities and an ODE analysis for the radial averages around the blow-up points) is then needed to show that no residual volume accumulates near the bubbling points. The analysis in [30] carries over with no substantial difficulties to closed manifolds, and one finds the following corollary.

Corollary 2.1 *Consider the problem* (1.5)*. Then for any compact set K of $\mathbb{R} \setminus 4\pi\mathbb{N}$ there exists a positive constant C_K such that all solutions of* (1.5) *are bounded in $C^{2,\alpha}(\Sigma)$ by C_K whenever $\rho \in K$.*

Remark 2.1 (i) When $h(x)$ is strictly positive on a compact manifold and is of class C^1, the local accumulation of mass for (1.5) near a blow-up point is exactly 4π. This was proved in [29] using moving-plane arguments, and allowed the author to prove that the Leray-Schauder degree of the equation is well defined for $\rho \notin 4\pi\mathbb{N}$. In [29], it was shown that the degree turns out to be always 1 when $\rho < 4\pi$. Using a more refined blow-up analysis and a Lyapunov-Schmidt reduction in [14], the Leray-Schauder degree of (1.5) was computed for general ρ's not belonging to $4\pi\mathbb{N}$ (see also [35] for a different approach).

(ii) It was proved instead in [15] that in bounded planar domains the mass accumulation can be a multiple of 4π strictly larger than 4π.

2.2 Compactness properties of (1.2)

We next turn to (1.2). It turns out that in this case there are different types of blow-ups, but it is still possible to obtain compactness results under rather neat assumptions. We first consider a sequence of solutions to a counterpart of (1.2), where the coefficients ρ_n are allowed to vary, namely

$$\begin{cases} -\Delta u_{1,n} = 2\rho_{1,n}\Big(\dfrac{h_1 \mathrm{e}^{u_{1,n}}}{\int_\Sigma h_1 \mathrm{e}^{u_{1,n}} \mathrm{d}V_g} - 1\Big) - \rho_{2,n}\Big(\frac{h_2 \mathrm{e}^{u_{2,n}}}{\int_\Sigma h_2 \mathrm{e}^{u_{2,n}} \mathrm{d}V_g} - 1\Big), \\ -\Delta u_{2,n} = 2\rho_{2,n}\Big(\dfrac{h_2 \mathrm{e}^{u_{2,n}}}{\int_\Sigma h_2 \mathrm{e}^{u_{2,n}} \mathrm{d}V_g} - 1\Big) - \rho_{1,n}\Big(\frac{h_1 \mathrm{e}^{u_{1,n}}}{\int_\Sigma h_1 \mathrm{e}^{u_{1,n}} \mathrm{d}V_g} - 1\Big). \end{cases} \tag{2.4}$$

Define the blow-up set

$$\widetilde{S} = \{x \in \Sigma : \exists x_n \to x \text{ such that } u_{i,n}(x_n) \to +\infty \text{ for some } i = 1, 2\}.$$

For a point $x \in \widetilde{S}$, we then define the local limit masses $A_i(x)$ as

$$A_i(x) = \lim_{r \to 0} \lim_n \int_{B_r(x)} \rho_{i,n}\Big(\frac{h_i \mathrm{e}^{u_{i,n}}}{\int_\Sigma h_i \mathrm{e}^{u_{i,n}} \mathrm{d}V_g} - 1\Big).$$

Then one has the following result (see also [32]).

Theorem 2.3 (see [25]) *Suppose that x is a blow-up point for* (2.4)*. Then only one among the following five possibilities may occur for $(A_1(x), A_2(x))$:*

$$(4\pi, 0), \quad (0, 4\pi), \quad (4\pi, 8\pi), \quad (8\pi, 4\pi), \quad (8\pi, 8\pi).$$

The reason for the restriction to these couples relies on the different blow-up rates of the two components. The first possibility (for the second one just exchange u_1 and u_2) corresponds to the case when the first component $u_{1,n}$ blows-up while the second does not. Then the situation is quite similar to the scalar case described before, with a blow-up profile given by (2.3).

The third possibility (again, exchange components for the fourth one) occurs when the blow-up rate of the first component is much faster than that of the second one. In this case, rescaling $u_{1,n}$ so that its maximal value becomes zero, the role of the second component $u_{2,n}$ will be irrelevant, so the limiting profile of $u_{1,n}$ will still be given by (2.3). On the other hand, looking at the equation satisfied by $u_{2,n}$ and rescaling so that $u_{2,n}$ has maximal height equal to zero, the term $\rho_{1,n}\big(\frac{h_1 e^{u_{1,n}}}{\int_\Sigma h_1 e^{u_{1,n}} dV_g} - 1\big)$ in the right-hand side of the equation will resemble a Dirac mass, with factor -4π. Therefore the profile of $u_{2,n}$ (after subtracting a suitable logarithmic function) will be given by the solution of the singular problem

$$-\Delta u = 2\overline{\rho}_2 e^u - 4\pi\delta_0 \quad \text{ in } \mathbb{R}^2. \tag{2.5}$$

Solutions of this singular equation were classified in [41], and these have the following expression:

$$\overline{U}(x) = -2\log(1 - 2|x|^2\cos 2(\theta-\theta_0)\tanh\xi + |x|^4) + C_{\zeta,\overline{\rho_2}}, \tag{2.6}$$

where $\xi > 0$ and where $\theta_0 \in [0, 2\pi)$. We notice that, compared to (2.3), the solution is not unique, it is not radial and it depends on the angular parameter θ_0. Indeed, if one considers the singular equation (2.5) but with a general weight $-\alpha$ in front of the Dirac mass, it turns out that solutions are always radial if α is not a positive multiple of 4π while, as already noticed in [12], there always exist non-radial solutions in the complementary case. However, as it happens for U_0 in (2.3), it turns out that $\overline{\rho}_2 \int_{\mathbb{R}^2} e^{\overline{U}} dx = 8\pi$ independently of $\overline{\rho}_2 > 0$. This fact yields the third alternative in Theorem 2.3.

Finally, the fifth alternative occurs when both components blow-up with the same rate. After scaling, the profile of $(u_{1,n}, u_{2,n})$ is given by the vectorial solution of the entire system

$$\begin{cases} -\Delta U_1 = 2\overline{\rho}_1 e^{U_1} - \overline{\rho}_2 e^{U_2}, \\ -\Delta U_2 = 2\overline{\rho}_2 e^{U_2} - \overline{\rho}_1 e^{U_1} \end{cases} \quad \text{in } \mathbb{R}^2. \tag{2.7}$$

Viewing (2.7) as a structure equation for holomorphic curves in $\mathbb{CP}^2$, such solutions were classified in [26]. These depend indeed on eight parameters, however it happens that one always has the quantization property

$$\overline{\rho}_1 \int_{\mathbb{R}^2} e^{U_1} dx = 8\pi, \quad \overline{\rho}_2 \int_{\mathbb{R}^2} e^{U_2} dx = 8\pi.$$

This yields the fifth possibility for the mass accumulation values in Theorem 2.3.

Using Green's representation formulas, in [6], it was proved that in case of blow-up at least one component u_i must accumulate at a finite number of points, and therefore the corresponding limiting parameter ρ_i must be quantized, according to Theorem 2.3. As a consequence one finds the following result.

Theorem 2.4 (see [6, 25]) *Consider the problem* (1.2). *Then for any compact set* K *of* $\mathbb{R} \setminus 4\pi\mathbb{N}$ *there exists a positive constant* C_K *such that all solutions of* (1.5) *are bounded in* $C^{2,\alpha}(\Sigma)$ *by* C_K *whenever* $\rho_i \in K$.

Remark 2.2 In [17], it was proved that there exist blowing-up solutions of (1.2) for which only one component concentrates near finitely-many points, while the other does not. Therefore, compactness holds true provided the couple (ρ_1, ρ_2) stays bounded away from the grid

$$\Lambda := \{(\rho_1, \rho_2) : \rho_i \in 4\pi\mathbb{N} \text{ for some } i = 1, 2\},$$

and not only from the squared lattice of points

$$\{(\rho_1, \rho_2) : \rho_i \in 4\pi\mathbb{N} \text{ for both } i = 1, 2\}.$$

As for (1.5), when $(\rho_1, \rho_2) \notin \Lambda$, the Leray-Schauder degree of (1.2) is well defined for $(\rho_1, \rho_2) \notin \Lambda$. Some degree-computations can be found in [31, 39] (and in [33] for other Liouville systems).

2.3 Applications to min-max theory

We will now show how to apply the previous results to deduce existence of solutions. Let us fix $(\rho_1, \rho_2) \notin \Lambda$, and let $\overline{\mathcal{K}} \subseteq \mathcal{K}$ be two compact sets. Let $\overline{\mathcal{F}} : \overline{\mathcal{K}} \to H^1(\Sigma) \times H^1(\Sigma)$ be a continuous map, and define the following class of continuous maps:

$$\mathfrak{F} := \{\mathcal{F} \in C(\mathcal{K}; H^1(\Sigma) \times H^1(\Sigma)) : \mathcal{F}|_{\overline{\mathcal{K}}} \equiv \overline{\mathcal{F}}\}.$$

Given this definition, one can introduce the corresponding min-max value

$$c_\rho := \inf_{\mathcal{F} \in \mathfrak{F}} \sup_{z \in \mathcal{K}} J_\rho(\mathcal{F}(z)).$$

As the Palais-Smale condition is not known yet for the functional J_ρ, one has to do some extra work in order to guarantee existence of critical points. We have the following proposition.

Proposition 2.1 *Suppose that there exists $t_0 > 0$ and $\beta_0 > 0$ such that $[\rho_1(1-t_0), \rho_1(1+t_0)] \times [\rho_2(1-t_0), \rho_2(1+t_0)] \subseteq \mathbb{R}^2 \setminus \Lambda$ and such that*

$$c_{(1+t)\rho} > \sup_{z \in \overline{\mathcal{K}}} J_\rho(\overline{\mathcal{F}}(z)) + \beta_0 \qquad \textit{for all } t \in [1-t_0, 1+t_0].$$

Then J_ρ admits a critical point at level c_ρ.

The proof of this result relies on both a monotonicity method introduced by Struwe in [42] (see also [34] for an alternative approach) and on the previous compactness results. First, one notices that since for $t' \geq t$

$$\frac{J_{t\rho}(u)}{t} - \frac{J_{t'\rho}(u)}{t'} = \left(\frac{1}{t} - \frac{1}{t'}\right) \int_\Sigma Q(u_1, u_2) \mathrm{d}V_g \geq 0, \quad u \in H^1(\Sigma) \times H^1(\Sigma),$$

so we clearly have that

$$\frac{\alpha_{t\rho}}{t} - \frac{\alpha_{t'\rho}}{t'} \geq 0.$$

This implies that the function $t \mapsto \frac{\alpha_{t\rho}}{t}$ is almost-everywhere differentiable. As in [19], one can prove that for the values of t for which $\frac{\alpha_{t\rho}}{t}$ is differentiable $J_{t\rho}$ has a bounded Palais-Smale sequence at level $\alpha_{t\rho}$. It can be shown that this sequence then converges to a critical point of $J_{t\rho}$. We will show it for the functional $I_{t\widetilde{\rho}}$ corresponding to the scalar equation (1.5) (with $\widetilde{\rho} \notin 4\pi\mathbb{N}$ and t close to 1). The vectorial counterpart requires only minor changes.

Consider a Palais-Smale sequence u_l for $I_{t\widetilde{\rho}}$ bounded in $H^1(\Sigma)$. The existence of a weak limit $u_0 \in H^1(\Sigma)$ follows from Theorem 2.2, as $t\widetilde{\rho} \notin 4\pi\mathbb{N}$ for t close to 1. Let us show that u_0 satisfies $I'_{t\widetilde{\rho}}(u_0) = 0$. For any function $v \in H^1(\Sigma)$ there holds

$$I'_{t\widetilde{\rho}}(u_0)[v] = I'_{t\widetilde{\rho}}(u_l)[v] + \int_\Sigma \nabla_g v \cdot \nabla_g (u_0 - u_l) \mathrm{d}V_g + t\widetilde{\rho}\Big(\frac{\int_\Sigma h e^{u_l} v \mathrm{d}V_g}{\int_\Sigma h e^{u_l} \mathrm{d}V_g} - \frac{\int_\Sigma h e^{u_0} v \mathrm{d}V_g}{\int_\Sigma h e^{u_0} \mathrm{d}V_g}\Big).$$

Since the first two terms in the right-hand side tend to zero by our assumptions, it is sufficient to check that $\int_\Sigma h e^{u_l} v \mathrm{d}V_g = \int_\Sigma h e^{u_0} v \mathrm{d}V_g + o(1)\|v\|_{H^1(\Sigma)}$ (to deal with the denominators just take $v \equiv 1$). In order to do this, we consider exponents $p, p', p'' > 1$ satisfying $\frac{1}{p} + \frac{1}{p'} + \frac{1}{p''} = 1$. Using Lagrange's formula, we obtain, for some function θ_l with range in $[0,1]$, $\mathrm{e}^{u_l} - \mathrm{e}^{u_0} = \mathrm{e}^{\theta_l u_l + (1-\theta_l) u_0}(u_l - u_0)$ almost everywhere in x. Then from some elementary inequalities (and the Moser-Trudinger inequality, that allows to control L^p norms of the exponentials), we find

$$\begin{aligned}\int_\Sigma (\mathrm{e}^{u_l} - \mathrm{e}^{u_0}) v \,\mathrm{d}V_g &\le C \int_\Sigma (\mathrm{e}^{u_l} + \mathrm{e}^{u_0}) |u_l - u_0| |v| \mathrm{d}V_g \\ &\le C[\|\mathrm{e}^{u_l}\|_{L^p(\Sigma)} + \|\mathrm{e}^{u_0}\|_{L^p(\Sigma)}] \|u_l - u_0\|_{L^{p'}(\Sigma)} \|v\|_{L^{p''}(\Sigma)} \\ &\le o(1) \|v\|_{L^{p''}(\Sigma)} = o(1) \|v\|_{H^1(\Sigma)},\end{aligned}$$

by (1.7), the boundedness of $(u_l)_l$ and the fact that $u_l \rightharpoonup u_0$ weakly in $H^1(\Sigma)$.

By the above reasoning, we proved that there exists $t_k \to 1$ for which the system

$$-\Delta u_{i,k} = \sum_{j=1}^{2} t_k \rho_j a_{ij} \Big(\frac{h_j \mathrm{e}^{u_{j,k}}}{\int_\Sigma h_j \mathrm{e}^{u_{j,k}} \mathrm{d}V_g} - 1 \Big), \quad i = 1, 2 \tag{2.8}$$

is solvable. Then using Theorem 2.4, which applies when both ρ_1, ρ_2 are not multiples of 4π, one obtains convergence of solutions and arrives to the desired conclusion.

3 Improved Inequalities

In this section, we collect some improved Moser-Trudinger type inequalities. These hold for functions satisfying certain requirements on the distribution of their exponentials. We will divide the discussion between cases in which the functions are macroscopically spread, and others that have scaling-invariant features.

3.1 Improved inequalities via macroscopic spreading

We present now the first improved inequality. Basically, if the mass of both $h_1 \mathrm{e}^{u_1}$ and $h_2 \mathrm{e}^{u_2}$ is spread respectively on at least $k+1$ and $l+1$ different sets, then the values of the ρ_i's for which one has coercivity increase by a factor $(k+1)$ and $(l+1)$ respectively.

We have first a couple of technical lemmas (see [7, Section 4] for details) that are useful for localizing the Moser-Trudinger inequality in Theorem 1.1.

Lemma 3.1 *Let $\delta > 0$ and $\Omega \Subset \widetilde{\Omega} \subset \Sigma$ be such that $d(\Omega, \partial\widetilde{\Omega}) \ge \delta$. Then, for any $\varepsilon > 0$ there exists $C = C(\varepsilon, \delta)$ such that for any $u = (u_1, u_2) \in H^1(\Sigma) \times H^1(\Sigma)$,*

$$\log \int_\Omega \mathrm{e}^{u_1 - \fint_\Omega u_1 \mathrm{d}V_g} \mathrm{d}V_g + \log \int_\Omega \mathrm{e}^{u_2 - \fint_\Omega u_2 \mathrm{d}V_g} \mathrm{d}V_g \le \frac{1}{4\pi} \int_{\widetilde{\Omega}} Q(u_1, u_2) \mathrm{d}V_g + \varepsilon \int_\Sigma Q(u_1, u_2) \mathrm{d}V_g + C.$$

Lemma 3.2 *Let $\delta > 0$, $\theta > 0$, $k, l \in \mathbb{N}$ with $k \ge l$, $f_i \in L^1(\Sigma)$ be non-negative functions with $\|f_i\|_{L^1(\Sigma)} = 1$ for $i = 1, 2$ and $\{\Omega_{1,i}, \Omega_{2,j}\}_{i \in \{0,\cdots,k\}, j \in \{0,\cdots,l\}} \subset \Sigma$ such that*

$$\begin{aligned} &d(\Omega_{1,i}, \Omega_{1,i'}) \ge \delta, \quad \forall\, i, i' \in \{0, \cdots, k\} \text{ with } i \ne i', \\ &d(\Omega_{2,j}, \Omega_{2,j'}) \ge \delta, \quad \forall\, j, j' \in \{0, \cdots, l\} \text{ with } j \ne j' \end{aligned}$$

and

$$\begin{aligned} &\int_{\Omega_{1,i}} f_1 \mathrm{d}V_g \ge \theta, \quad \forall\, i \in \{0, \cdots, k\}, \\ &\int_{\Omega_{2,j}} f_2 \mathrm{d}V_g \ge \theta, \quad \forall\, j \in \{0, \cdots, l\}. \end{aligned}$$

Then, there exist $\overline{\delta} > 0$, $\overline{\theta} > 0$, *independent of* f_i, *and* $\{\Omega_n\}_{n=1}^{k} \subset \Sigma$ *such that*

$$d(\Omega_n, \Omega_{n'}) \geq \overline{\delta}, \quad \forall\, n, n' \in \{0, \cdots, k\} \text{ with } n \neq n'$$

and

$$\begin{aligned} |\Omega_n| &\geq \overline{\theta}, \quad \forall\, n \in \{0, \cdots, k\}, \\ \int_{\Omega_n} f_1 \mathrm{d}V_g &\geq \overline{\theta}, \quad \forall\, n \in \{0, \cdots, k\}, \\ \int_{\Omega_n} f_2 \mathrm{d}V_g &\geq \overline{\theta}, \quad \forall\, n \in \{0, \cdots, l\}. \end{aligned}$$

We then have the following result: It says qualitatively that the more the components (u_1, u_2) of the system are spread over Σ, the more effectively $Q(u_1, u_2)$ controls the exponential integrals.

Proposition 3.1 (see [7]) *Let* $\delta > 0$, $\theta > 0$, $k, l \in \mathbb{N}$ *and* $\{\Omega_{1,i}, \Omega_{2,j}\}_{i \in \{0,\cdots,k\}, j \in \{0,\cdots,l\}} \subset \Sigma$ *be such that*

$$\begin{aligned} d(\Omega_{1,i}, \Omega_{1,i'}) &\geq \delta, \quad \forall\, i, i' \in \{0, \cdots, k\} \text{ with } i \neq i', \\ d(\Omega_{2,j}, \Omega_{2,j'}) &\geq \delta, \quad \forall\, j, j' \in \{0, \cdots, l\} \text{ with } j \neq j'. \end{aligned}$$

Then, for any $\varepsilon > 0$, *there exists* $C = C(\varepsilon, \delta, \theta, k, l, \Sigma)$, *such that any* $u = (u_1, u_2) \in H^1(\Sigma) \times H^1(\Sigma)$ *satisfying*

$$\begin{aligned} \int_{\Omega_{1,i}} h_1 \mathrm{e}^{u_1} \mathrm{d}V_g &\geq \theta \int_{\Sigma} h_1 \mathrm{e}^{u_1} \mathrm{d}V_g, \quad \forall\, i \in \{0, \cdots, k\}, \\ \int_{\Omega_{2,j}} h_2 \mathrm{e}^{u_2} \mathrm{d}V_g &\geq \theta \int_{\Sigma} h_2 \mathrm{e}^{u_2} \mathrm{d}V_g, \quad \forall\, j \in \{0, \cdots, l\} \end{aligned}$$

verifies

$$(k+1) \log \int_{\Sigma} h_1 \mathrm{e}^{u_1 - \overline{u_1}} \mathrm{d}V_g + (l+1) \log \int_{\Sigma} h_2 \mathrm{e}^{u_2 - \overline{u_2}} \mathrm{d}V_g \leq \frac{1+\varepsilon}{4\pi} \int_{\Sigma} Q(u_1, u_2) \mathrm{d}V_g + C.$$

Proof In the proof, we assume that $\overline{u_1} = \overline{u_2} = 0$. After relabelling the indexes, we can suppose $k \geq l$ and apply Lemma 3.2 with $f_i = \frac{h_i \mathrm{e}^{u_i}}{\int_{\Sigma} h_i \mathrm{e}^{u_i} \mathrm{d}V_g}$ to get $\{\Omega_j\}_{j=0}^{k} \subset \Sigma$ with

$$d(\Omega_i, \Omega_j) \geq \overline{\delta}, \quad \forall\, i, j \in \{0, \cdots, k\} \ \text{ with } i \neq j$$

and

$$\begin{aligned} \int_{\Omega_i} h_1 \mathrm{e}^{u_1} \mathrm{d}V_g &\geq \overline{\theta} \int_{\Sigma} h_1 \mathrm{e}^{u_1} \mathrm{d}V_g, \quad \forall\, i \in \{0, \cdots, k\}, \\ \int_{\Omega_j} h_2 \mathrm{e}^{u_2} \mathrm{d}V_g &\geq \overline{\theta} \int_{\Sigma} h_2 \mathrm{e}^{u_2} \mathrm{d}V_g, \quad \forall\, j \in \{0, \cdots, l\}. \end{aligned}$$

Notice that

$$\log \int_{\Sigma} h_i \mathrm{e}^{u_i} \mathrm{d}V_g = \fint_{\widetilde{\Omega}_j} u_i \mathrm{d}V_g + \log \int_{\Sigma} h_i \mathrm{e}^{u_i - \fint_{\widetilde{\Omega}_j} u_i \mathrm{d}V_g} \mathrm{d}V_g, \quad i = 1, 2.$$

The average on $\widetilde{\Omega}_j$ can be estimated by Poincaré's inequality as follows:

$$\fint_{\widetilde{\Omega}_j} u_i \mathrm{d}V_g \le \frac{1}{|\widetilde{\Omega}_j|}\int_\Sigma |u_i|\mathrm{d}V_g \le C\Big(\int_\Sigma |\nabla u_i|^2 \mathrm{d}V_g\Big)^{\frac{1}{2}} \le C + \varepsilon \int_\Sigma |\nabla u_i|^2 \mathrm{d}V_g, \quad i = 1, 2. \tag{3.1}$$

We now apply, for any $j \in \{0, \cdots, k\}$, Lemma 3.1 with $\Omega = \Omega_j$ and $\widetilde{\Omega} = \widetilde{\Omega}_j := \big\{x \in \Sigma : d(x, \Omega_j) < \frac{\overline{\delta}}{2}\big\}$. For $j \in \{0, \cdots, l\}$, we get

$$\begin{aligned}
&\log \int_\Sigma h_1 \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g + \log \int_\Sigma h_2 \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g \\
&\le 2 \log \frac{1}{\overline{\theta}} + \log \int_{\Omega_j} h_1 \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g + \log \int_{\Omega_j} h_2 \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g \\
&\le C + \log \int_{\Omega_j} \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g + \log \int_{\Omega_j} \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g \\
&\le C + \frac{1}{4\pi} \int_{\widetilde{\Omega}_j} Q(u_1, u_2) \mathrm{d}V_g + \varepsilon \int_\Sigma Q(u_1, u_2) \mathrm{d}V_g, \quad j = 1, \cdots, l.
\end{aligned} \tag{3.2}$$

For $j \in \{l+1, \cdots, k\}$, we have

$$\begin{aligned}
\log \int_\Sigma h_1 \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g &\le \log \frac{1}{\overline{\theta}} + \|h_1\|_{L^\infty(\Sigma)} + \log \int_{\Omega_j} \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g \\
&\le C - \log \int_{\Omega_j} \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g + \frac{1}{4\pi} \int_{\widetilde{\Omega}_j} Q(u_1, u_2) \mathrm{d}V_g \\
&\quad + \varepsilon \int_\Sigma Q(u_1, u_2) \mathrm{d}V_g.
\end{aligned} \tag{3.3}$$

The exponential term on the second component can be estimated by using Jensen's inequality as follows:

$$\log \int_{\Omega_j} \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g = \log |\Omega_j| + \log \fint_{\Omega_j} \mathrm{e}^{u_2 - \fint_{\widetilde{\Omega}_j} u_2 \mathrm{d}V_g} \mathrm{d}V_g \ge \log |\Omega_j| \ge -C. \tag{3.4}$$

Putting together (3.3) and (3.4), we have

$$\begin{aligned}
&\log \int_\Sigma h_1 \mathrm{e}^{u_1 - \fint_{\widetilde{\Omega}_j} u_1 \mathrm{d}V_g} \mathrm{d}V_g \\
&\le \frac{1}{4\pi} \int_{\widetilde{\Omega}_j} Q(u_1, u_2) \mathrm{d}V_g + \varepsilon \int_\Sigma Q(u_1, u_2) \mathrm{d}V_g + C, \quad j = l+1, \cdots, k.
\end{aligned} \tag{3.5}$$

Summing over all $j \in \{0, \cdots, k\}$ and taking into account (3.2) and (3.5), we obtain the result, renaming ε appropriately.

3.2 Scaling-invariant improved inequalities

We next introduce two new improved inequalities that have the feature of being scaling-invariant. As we already remarked, this means that can be applied to functions which might be indefinitely concentrated near a single point, differently from the previous proposition.

We begin by considering the family of functions

$$A = \Big\{ f \in L^1(\Sigma) : f > 0 \text{ a.e. and} \int_\Sigma f \mathrm{d}V_g = 1 \Big\}.$$

First we have the following result about functions in L^1 that are sufficiently concentrated near a single point.

Proposition 3.2 *Fix $R > 1$. Then there exists $\delta = \delta(R) > 0$ and a continuous map*

$$\psi : A \cap \{\mathbf{d}(\cdot, \Sigma_1) < \delta(R)\} \to \Sigma \times \mathbb{R}_+, \quad \psi(f) = (\beta, \sigma),$$

satisfying the following property: Given $f \in A$ there exists $p \in \Sigma$ such that

(a) $d(p, \beta) \leq C'\sigma$ *for* $C' = \max\{3R + 1, \delta^{-1}\mathrm{diam}(\Sigma)\}$,

(b) *there exists $\tau > 0$ depending only on R and Σ such that*

$$\int_{B_p(\sigma)} f \, \mathrm{d}V_g > \tau, \qquad \int_{B_p(R\sigma)^c} f \, \mathrm{d}V_g > \tau.$$

Proof We only sketch the main arguments, referring to [38] for full details. Take $R_0 = 3R$, and let $\sigma : \Sigma \times A \to (0, +\infty)$ (well defined and continuous) be such that

$$\int_{B_x(\sigma(x,f))} f \, \mathrm{d}V_g = \int_{B_x(R_0\sigma(x,f))^c} f \, \mathrm{d}V_g. \tag{3.6}$$

We notice that σ satisfies

$$d(x, y) \leq R_0 \max\{\sigma(x, f), \sigma(y, f)\} + \min\{\sigma(x, f), \sigma(y, f)\}. \tag{3.7}$$

In fact, if this were not true, we would have $B_x(R_0\sigma(x, f)) \cap B_y(\sigma(y, f) + \varepsilon) = \emptyset$ for some $\varepsilon > 0$. Also, $B_y(R_0\sigma(y, f))$ cannot coincide with Σ, so $A_y(\sigma(y, f), \sigma(y, f) + \varepsilon)$ ($A_y(r_1, r_2)$ stands for the open annulus centered at y with radii r_1, r_2) is non-empty and open. This implies that

$$\int_{B_x(\sigma(x,f))} f \, \mathrm{d}V_g = \int_{B_x(R_0\sigma(x,f))^c} f \, \mathrm{d}V_g \geq \int_{B_y(\sigma(y,f)+\varepsilon)} f \, \mathrm{d}V_g > \int_{B_y(\sigma(y,f))} f \, \mathrm{d}V_g.$$

By interchanging x and y, we also obtain the opposite inequality, which proves (3.7).

Next, setting

$$T : \Sigma \times A \to \mathbb{R}, \quad T(x, f) = \int_{B_x(\sigma(x,f))} f \mathrm{d}V_g,$$

we make the following claim.

Claim 3.1 If $x_0 \in \Sigma$ satisfies $T(x_0, f) = \max_{y \in \Sigma} T(y, f)$, then $\sigma(x_0, f) < 3\sigma(x, f)$ for any other $x \neq x_0$.

To see this, fix $x \in \Sigma$ and $\varepsilon > 0$. First, reasoning as above we find that $B_x(R_0\sigma(x, f) + \varepsilon) \cap B_{x_0}(\sigma(x_0, f)) \neq \emptyset$, and similarly that $B_x(R_0\sigma(x, f) + \varepsilon)$ cannot be contained in $B_{x_0}(R_0\sigma(x_0, f))$. From the triangular inequality, one has

$$2(R_0\sigma(x, f) + \varepsilon) > (R_0 - 1)\sigma(x_0, f),$$

so by the arbitrariness of ε we get that $\sigma(x, f) \geq \frac{R_0 - 1}{2R_0}\sigma(x_0, f)$. The claim follows from the fact that $R_0 > 3$.

Using a covering argument, one also has that there exists a $\tau > 0$ (independent of f) such that

$$\max_{x \in \Sigma} T(x, f) > \tau > 0 \quad \text{for all } f \in A. \tag{3.8}$$

Let us now fix $x_0 \in \Sigma$ such that $T(x_0, f) = \max_{x\in\Sigma} T(x, f)$. By the above claim, for any $x \in A_{x_0}(\sigma(x_0, f), R\sigma(x_0, f))$, one has

$$\int_{B_x \frac{\sigma(x_0,f)}{3}} f \, \mathrm{d}V_g \leq \int_{B_x(\sigma(x,f))} f \, \mathrm{d}V_g \leq T(x_0, f).$$

Taking a finite covering of the form

$$A_{x_0}(\sigma(x_0, f), R\sigma(x_0, f)) \subset \bigcup_{i=1}^{k} B_{x_i} \frac{\sigma(x_0, f)}{3}$$

(where k can be chosen depending only on Σ and R), we find

$$\begin{aligned} 1 = \int_\Sigma f \, \mathrm{d}V_g &\leq \int_{B_{x_0}(\sigma(x_0,f))} f \, \mathrm{d}V_g + \int_{B_{x_0}(R\sigma(x_0,f))^c} f \, \mathrm{d}V_g + \sum_{i=1}^{k} \int_{B_{x_i} \frac{\sigma(x_0,f)}{3}} f \, \mathrm{d}V_g \\ &\leq (k+2)T(x_0, f). \end{aligned}$$

Considering the continuous function

$$\sigma : A \to \mathbb{R}, \quad \sigma(f) = 3 \min\{\sigma(x, f) : x \in \Sigma\},$$

and given τ as in (3.8), define

$$\mathcal{S}(f) = \{x \in \Sigma : T(x, f) > \tau, \ \sigma(x, f) < \sigma(f)\}. \tag{3.9}$$

Claim 3.1 and (3.8) imply that if $x_0 \in \Sigma$ maximizes $T(x, f)$, then $x_0 \in \mathcal{S}(f)$. Hence for any $f \in A$ $\mathcal{S}(f)$ is non-empty and open. Moreover, (3.7) implies

$$\mathrm{diam}(\mathcal{S}(f)) \leq (R_0 + 1)\sigma(f). \tag{3.10}$$

Embedding Σ in $\mathbb{R}^3$ and identifying it with its image, we define the center of mass

$$\eta(f) = \frac{\int_\Sigma (T(x, f) - \tau)^+ (\sigma(f) - \sigma(x, f))^+ x \ \mathrm{d}V_g}{\int_\Sigma (T(x, f) - \tau)^+ (\sigma(f) - \sigma(x, f))^+ \mathrm{d}V_g} \in \mathbb{R}^N.$$

For $\delta > 0$ small, let P be an orthogonal projection from a δ-neighbourhood of Σ onto the surface, and define

$$\beta : \{f \in A : \ \sigma(f) \leq \delta\} \to \Sigma, \ \ \beta(f) = P \circ \eta(f).$$

To conclude the proof, we check that $\psi(f) = (\beta(f), \sigma(f))$ satisfies the desired condition. If $\sigma(f) \leq \delta$, then $d(\beta(f), \mathcal{S}(f)) < (R_0 + 1)\sigma(f)$. Taking $p \in \mathcal{S}(f)$, recalling that $R_0 = 3R$ and that $\sigma(f) \leq 3\sigma(x, f) < 3\sigma(f)$ for any $x \in \mathcal{S}(f)$, we then deduce both (a) and (b).

The next result provides a lower bound on the functional in terms of the function ψ. Its proof relies on using Kelvin inversions, which preserve the integral of the quadratic form Q.

Proposition 3.3 (see [37]) *Given any $\varepsilon > 0$, there exist $R = R(\varepsilon) > 1$ and ψ as in Proposition* 3.2 *for which, if*

$$\psi\Big(\frac{\mathrm{e}^{u_1}}{\int_\Sigma \mathrm{e}^{u_1} \mathrm{d}V_g}\Big) = \psi\Big(\frac{\mathrm{e}^{u_2}}{\int_\Sigma \mathrm{e}^{u_2} \mathrm{d}V_g}\Big),$$

then there exists $C = C(\varepsilon)$ such that

$$(1 + \varepsilon) \int_\Sigma Q(u_1, u_2) \, \mathrm{d}V_g \geq 8\pi \Big(\log \int_\Sigma \mathrm{e}^{u_1 - \overline{u}_1} \mathrm{d}V_g + \log \int_\Sigma \mathrm{e}^{u_2 - \overline{u}_2} \mathrm{d}V_g \Big) + C.$$

The previous result states roughly that if the two components have the same scale of concentration and near the same point, then the Moser-Trudinger constant improves. The next proposition applies instead to the case in which one component (u_1) is much more concentrated than the other.

Proposition 3.4 (see [24]) *Let $r > 0$, $\gamma_0 > 0$ and $\tau_0 > 0$. For any $\varepsilon > 0$, there exists $C = C(\varepsilon, r, \tau_0, \gamma_0)$ such that, if for some $\sigma \in \left(0, \frac{r}{C^2}\right)$ and $z \in \Sigma$, it holds that*

$$\frac{\int_{B_{\frac{\sigma}{2}}(z)} \mathrm{e}^{u_1}\,\mathrm{d}V_g}{\int_\Sigma \mathrm{e}^{u_1}\,\mathrm{d}V_g} > \gamma_0, \qquad \frac{\int_{A_z(C\sigma, \frac{r}{C})} \mathrm{e}^{u_2}\,\mathrm{d}V_g}{\int_\Sigma \mathrm{e}^{u_2}\,\mathrm{d}V_g} > \gamma_0 \tag{3.11}$$

and

$$\sup_{y \in A_z(C\sigma, \frac{r}{C})} \frac{\int_{B_{\tau_0 d(y,z)}(y)} \mathrm{e}^{u_2}\,\mathrm{d}V_g}{\int_{A_z(C\sigma, \frac{r}{C})} \mathrm{e}^{u_2}\,\mathrm{d}V_g} < 1 - \tau_0, \tag{3.12}$$

then

$$4\pi \log \int_\Sigma \mathrm{e}^{u_1 - \overline{u}_1}\,\mathrm{d}V_g + 8\pi \log \int_\Sigma \mathrm{e}^{u_2 - \overline{u}_2}\,\mathrm{d}V_g \le \int_{B_r(z)} Q(u_1, u_2)\,\mathrm{d}V_g + \varepsilon \int_\Sigma Q(u_1, u_2)\,\mathrm{d}V_g + C.$$

The requirement in (3.12) means qualitatively that e^{u_2} is well distributed around the concentration point (with smaller scale) of e^{u_1}. The result is inspired from a similar situation regarding the singular Liouville problem, whose relation to (1.2) was discussed in Section 1. In this case, one has the following improved inequality.

Proposition 3.5 (see [4]) *Let $p \in \Sigma$ and $r > 0$, $\tau_0 > 0$. Then, for any $\varepsilon > 0$, there exists $C = C(\varepsilon, r)$ such that*

$$\log \int_{B_r(p)} d(x,p)^2 \mathrm{e}^v\,\mathrm{d}V_g \le \frac{1+\varepsilon}{32\pi} \int_{B_r(p)} |\nabla v|^2\,\mathrm{d}V_g + C$$

for every function $v \in H_0^1(B_r(p))$ such that

$$\sup_{y \in B_r(p);\, y \ne p} \frac{\int_{B_{\tau_0 d(y,p)}(y)} d(x,p)^2 \mathrm{e}^v\,\mathrm{d}V_g}{\int_{B_r(p)} d(x,p)^2 \mathrm{e}^v\,\mathrm{d}V_g} < 1 - \tau_0.$$

In [4], the assumption in the last proposition was also expressed in terms of the angular moments of the function $d(x,p)^2\mathrm{e}^v$ around the singular point p.

4 Weighted Barycentric Sets and Topological Join

In this section, we characterize the low-energy levels of J_ρ in supercritical regimes. We show first that the improved inequality in Proposition 3.1 leads naturally to consider barycentric set of Σ and their topological join. Then, the scaling-invariant improved inequalities in Propositions 3.3–3.4 are used to provide further properties of low-energy functions.

4.1 General description of low energy levels

We now state a technical result that gives sufficient conditions to apply Proposition 3.1. Its proof relies on a covering argument.

Lemma 4.1 (see [36, 20]) *Let $f \in L^1(\Sigma)$ be a non-negative function with $\|f\|_{L^1(\Sigma)} = 1$, and let $m \in \mathbb{N}$ be such that there exist $\varepsilon > 0,\ s > 0$ with*

$$\int_{\bigcup_{j=0}^m B_s(x_j)} f \mathrm{d}V_g < 1 - \varepsilon, \quad \forall \{x_j\}_{j=0}^m \subset \Sigma.$$

Then there exist $\overline{\varepsilon} > 0,\ \overline{s} > 0$, not depending on f, and $\{\overline{x}_j\}_{j=1}^m \subset \Sigma$ satisfying

$$\int_{B_{\overline{s}}(\overline{x}_j)} f \mathrm{d}V_g > \overline{\varepsilon}, \quad \forall\, j \in \{1, \cdots, m\},$$
$$B_{2\overline{s}}(\overline{x}_i) \cap B_{2\overline{s}}(\overline{x}_j) = \emptyset, \quad \forall\, i, j \in \{1, \cdots, m\}, \quad i \neq j.$$

Applying this result to both $h_1(x)\mathrm{e}^{u_1}$ and $h_2(x)\mathrm{e}^{u_2}$ (once normalized in L^1), jointly with Proposition 3.1, we obtain the following concentration alternative for the exponential functions.

Lemma 4.2 *Suppose $\rho_1 \in (4k\pi, 4(k+1)\pi)$ and $\rho_2 \in (4l\pi, 4(l+1)\pi)$. Then, for any $\varepsilon > 0$, $s > 0$, there exists $L = L(\varepsilon, s) > 0$ such that for any $u \in J_\rho^{-L}$ there are either some $\{x_i\}_{i=1}^k \subset \Sigma$ such that*

$$\frac{\int_{\bigcup_{i=1}^k B_s(x_i)} h_1 \mathrm{e}^{u_1} \mathrm{d}V_g}{\int_\Sigma h_1 \mathrm{e}^{u_1} \mathrm{d}V_g} \geq 1 - \varepsilon$$

or some $\{y_j\}_{j=1}^l \subset \Sigma$ such that

$$\frac{\int_{\bigcup_{j=1}^l B_s(y_j)} h_2 \mathrm{e}^{u_2} \mathrm{d}V_g}{\int_\Sigma h_2 \mathrm{e}^{u_2} \mathrm{d}V_g} \geq 1 - \varepsilon.$$

An immediate consequence of the previous lemma is that at least one of the two $h_i\mathrm{e}^{u_i}$'s (once normalized in L^1) has to be distributionally close respectively to the sets of k-barycenters or l-barycenters of Σ.

Corollary 4.1 *Suppose $\rho_1 \in (4k\pi, 4(k+1)\pi)$ and $\rho_2 \in (4l\pi, 4(l+1)\pi)$. Then, for any $\varepsilon > 0$, there exists $L > 0$ such that any $u \in J_\rho^{-L}$ verifies either*

$$\mathbf{d}\Big(\frac{h_1 \mathrm{e}^{u_1}}{\int_\Sigma h_1 \mathrm{e}^{u_1} \mathrm{d}V_g}, \Sigma_k\Big) < \varepsilon \quad \text{or} \quad \mathbf{d}\Big(\frac{h_2 \mathrm{e}^{u_2}}{\int_\Sigma h_2 \mathrm{e}^{u_2} \mathrm{d}V_g}, \Sigma_l\Big) < \varepsilon.$$

We can now see the role of the topological join of the barycentric sets of Σ.

Proposition 4.1 *Suppose $\rho_1 \in (4k\pi, 4(k+1)\pi)$, $\rho_2 \in (4l\pi, 4(l+1)\pi)$, and let Φ_λ be as in* (5.2). *Then for L sufficiently large, there exists a natural continuous map*

$$\Psi : J_\rho^{-L} \to \Sigma_k * \Sigma_l$$

from low-energy levels of J_ρ into the topological join of Σ_k and Σ_l.

By natural, we mean that we are able to construct a sort of right-inverse of this map (see the next section for details).

Proof of Proposition 4.1 It was proved in [20–21] that if $m \in \mathbb{N}$, then there exists a retraction ψ_m from a small neighbourhood of Σ_m (with respect to the distance $\mathbf{d}$ defined after (1.10)) onto Σ_m.

By Corollary 4.1, we know that either $\psi_k\big(\frac{h_1 e^{u_1}}{\int_\Sigma h_1 e^{u_1} dV_g}\big)$ or $\psi_l\big(\frac{h_2 e^{u_2}}{\int_\Sigma h_2 e^{u_2} dV_g}\big)$ is well defined (or both), since either $\mathbf{d}\big(\frac{h_1 e^{u_1}}{\int_\Sigma h_1 e^{u_1} dV_g}, \Sigma_k\big) < \varepsilon$ or $\mathbf{d}\big(\frac{h_2 e^{u_2}}{\int_\Sigma h_2 e^{u_2} dV_g}, \Sigma_l\big) < \varepsilon$ (or both).

We then set

$$d_1 = \mathbf{d}\Big(\frac{h_1 e^{u_1}}{\int_\Sigma h_1 e^{u_1} dV_g}, \Sigma_k\Big), \quad d_2 = \mathbf{d}\Big(\frac{h_2 e^{u_2}}{\int_\Sigma h_2 e^{u_2} dV_g}, \Sigma_l\Big),$$

and consider a function $\widetilde{s} = \widetilde{s}(d_1, d_2)$ with the expression

$$\widetilde{s}(d_1, d_2) = \mathfrak{f}\Big(\frac{d_1}{d_1 + d_2}\Big), \tag{4.1}$$

where $\mathfrak{f}$ is such that

$$\mathfrak{f}(z) = \begin{cases} 0, & \text{if } z \in \Big[0, \dfrac{1}{4}\Big], \\ 2z - \frac{1}{2}, & \text{if } z \in \Big(\dfrac{1}{4}, \dfrac{3}{4}\Big), \\ 1, & \text{if } z \in \Big[\dfrac{3}{4}, 1\Big]. \end{cases} \tag{4.2}$$

The desired map is then defined by

$$\Psi(u_1, u_2) = \Big(\psi_k\Big(\frac{h_1 e^{u_1}}{\int_\Sigma h_1 e^{u_1} dV_g}\Big), \psi_l\Big(\frac{h_2 e^{u_2}}{\int_\Sigma h_2 e^{u_2} dV_g}\Big), \widetilde{s}\Big), \tag{4.3}$$

where we are using the notation in (1.9).

4.2 Surfaces with positive genus

In this section, we consider the case of positive genus, where the map from Proposition 4.1 will be specialized. We begin with an easy topological result, whose proof is evident from the picture below.

Lemma 4.3 *Let Σ be a compact surface with positive genus. Then, there exist two simple closed curves $\gamma_1, \gamma_2 \subseteq \Sigma$ satisfying*

(1) *γ_1, γ_2 do not intersect each other,*

(2) *there exist global retractions $\Pi_i : \Sigma \to \gamma_i$, $i = 1, 2$.*

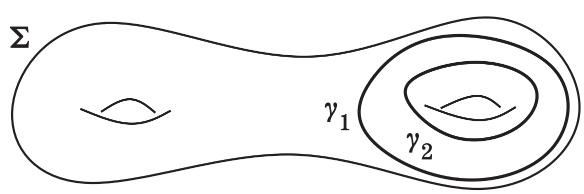

Figure 1 A compact surface with positive genus Σ.

Consider the global retractions $\Pi_1 : \Sigma \to \gamma_1$ and $\Pi_2 : \Sigma \to \gamma_2$ given in Lemma 4.3. Acting by push-forward, any probability measure on Σ is sent by $(\Pi_1)_*$ (resp., by $(\Pi_2)_*$) into a probability measure on γ_1 (resp., on γ_2). In this case, Proposition 4.1 has the following variant, that is quite useful for our purposes.

Proposition 4.2 *Suppose $\rho_1 \in (4k\pi, 4(k+1)\pi)$, $\rho_2 \in (4l\pi, 4(l+1)\pi)$. Then for L sufficiently large, there exists a natural continuous map*

$$\widetilde{\Psi} : J_\rho^{-L} \to (\gamma_1)_k * (\gamma_2)_l$$

from low-energy levels of J_ρ into the topological join of $(\gamma_1)_k$ and $(\gamma_2)_l$.

Remark 4.1 Since each γ_i is homeomorphic to S^1, it follows from [9, Proposition 3.2] that $(\gamma_1)_k$ is homeomorphic to S^{2k-1} and $(\gamma_2)_l$ to S^{2l-1} (in [27], it was proved previously a homotopy equivalence). As it is well-known, the topological join $S^m * S^n$ is homeomorphic to S^{m+n+1} (see, for example, [23]), and therefore $(\gamma_1)_k * (\gamma_2)_l$ is homeomorphic to the sphere $S^{2k+2l-1}$.

4.3 Constraints from scaling-invariant inequalities

In this subsection, we make use of the scaling-invariant improved inequalities from the previous section in order to find some constraints on the maps from low-energy levels into the topological join of the barycentric sets.

We first consider the case $(\rho_1, \rho_2) \in (4\pi, 8\pi)$. We perform a construction similar to (4.3), but taking the scales of concentration of the u_i's (as defined in Proposition 3.2) into account. Notice that the scale σ is only defined in a $\delta(R)$ (the choice of R will be made before Proposition 4.3) neighbourhood (with respect to the distance $\mathbf{d}$). To extend this scale to arbitrary functions, we set

$$\widehat{\sigma}_1 = \inf \left\{ \sigma(f) : \mathbf{d}(f, \Sigma_1) \le \frac{1}{2}\delta_R \right\},$$

and then

$$\overline{\sigma}(u_i) = \min \left\{ \widehat{\sigma}_1, \sigma\Big(\frac{h_i e^{u_i}}{\int_\Sigma h_i e^{u_i} dV_g} \Big) \right\}$$

with the convention of choosing $\widehat{\sigma}_1$ whenever $\sigma\big(\frac{h_i e^{u_i}}{\int_\Sigma h_i e^{u_i} dV_g}\big)$ is not well defined.

If $\mathfrak{f}$ is as in (4.1), we define a modified map $\widehat{\Psi}$ as

$$\widehat{\Psi}(u_1, u_2) = (\beta(u_1), \beta(u_2), \widetilde{s}(\overline{\sigma}(u_1), \overline{\sigma}(u_2))). \tag{4.4}$$

By means of Proposition 3.3, we then deduce the following result, which imposes some constraint on the map into the topological join $\Sigma_1 * \Sigma_1 \simeq \Sigma * \Sigma$ (the number ε in Proposition 3.3 will be taken sufficiently small, and R in Proposition 3.2 taken as $R(\varepsilon)$).

Proposition 4.3 *For $(\rho_1, \rho_2) \in (4\pi, 8\pi)$, let $\widehat{\Psi}$ be as in (4.4). Then for L sufficiently large $\widehat{\Psi}$ sends J_ρ^{-L} into $(\Sigma * \Sigma) \setminus \widehat{S}$, where*

$$\widehat{S} := \left\{ (x, x, \frac{1}{2}) \ : \ x \in \Sigma \right\} \subseteq \Sigma * \Sigma.$$

When $k > 1$ in Theorem 1.4, for $\delta > 0$ small, we define the set

$$\check{S} = \left\{ \Big(\nu, \delta_z, \frac{1}{2}\Big) \in \Sigma_k * \Sigma_1 : \nu = \sum_{i=1}^k t_i \delta_{x_i};\ d(x_i, x_j) \ge \delta,\ \forall\, i \ne j, \right.$$
$$\left. \delta \le t_i \le 1 - \delta,\ \forall\, i;\ z \in \operatorname{supp}(\nu) \right\}. \tag{4.5}$$

The counterpart of the above proposition becomes the following one.

Proposition 4.4 *Let ρ_1, ρ_2 be as in Theorem* 1.4*, with $k \geq 2$. Let $\check{S}$ be as in* (4.5) *and let $Y = (\Sigma_k * \Sigma_1) \setminus \check{S}$. Then, for $L > 0$ large there exists a continuous map $\check{\Psi}$ from the low sublevels J_ρ^{-L} into the set Y.*

We limit ourselves to give just few ideas beyond this result, referring to [24, Section 3] for the details. Under the assumptions of Theorem 1.4, when $k \geq 1$, we have that for low values of $J_\rho(u)$ either e^{u_1} is concentrated near at most k points of Σ or e^{u_2} is concentrated near a single point. From the construction in the proof of Proposition 4.1, the join parameter is chosen depending on the **d**-distances of the exponentials from Σ_k and $\Sigma_1 \simeq \Sigma$.

In a situation when both components are concentrated, we would also like to take into account the relative scales of the two components, as it was done in (4.4). For u_2, which is concentrated near a single point, a natural scale to use is the function β from Proposition 3.2. For u_1, which might be concentrated near multiple points (recall that now $k \geq 2$), there is a way to localize this quantity near each peak, and to choose the one for the peak closest to that of u_2. The latter definition might sound ambiguous because of possible multiple choices, but there is a rigorous way to define a scale of u_1 near the peak of u_2 by an averaging process. The choice of the join parameter should then take also into account the ratios of the two scales (absolute for u_2 and local near the peak of u_2 for u_1).

Now, two competing effects might take place in determining the join parameter. On the one hand, a small local scale of u_1 relative to that of u_2 would tend the join parameter to approach 0. On the other hand, having $\mathbf{d}(e^{u_1}, \Sigma_k)$ not that small would make the join parameter approach 1. This is precisely the situation in the assumption of Proposition 3.4. u_1 has a peak sharper than that of u_2 (and near the same point), but at the same time starts to split (at a macroscopic level) into $k+1$ regions (see Lemma 4.1). One can then combine (a localized version of) Proposition 3.4 and Proposition 3.1 to get a lower bound on the energy.

5 Proofs of the Theorems

In this section, we sketch the proofs of Theorems 1.3–1.4. We first construct suitable test functions modelled on (proper subsets of the) the topological joins, and then introduce variational min-max schemes in order to find solutions as saddle points of the Euler-Lagrange energy.

5.1 Test functions

We first consider the case of positive genus. For $\rho_1 \in (4k\pi, 4(k+1)\pi)$ and $\rho_2 \in (4l\pi, 4(l+1)\pi)$, we wish to build a family of test functions modelled on the topological join $(\gamma_1)_k * (\gamma_2)_l$, the barycentric sets of the curves γ_1, γ_2 (see Lemma 4.3).

Let $\zeta = (\sigma_1, \sigma_2, r) \in (\gamma_1)_k * (\gamma_2)_l$, where

$$\sigma_1 := \sum_{i=1}^{k} t_i \delta_{x_i} \in (\gamma_1)_k \quad \text{and} \quad \sigma_2 := \sum_{j=1}^{l} s_j \delta_{y_j} \in (\gamma_2)_l.$$

Our goal is to define a test function modelled on any $\zeta \in (\gamma_1)_k * (\gamma_2)_l$, depending on a positive parameter λ and belonging to low sub-levels of J_ρ for large λ, that is, to find a map

$$\Phi_\lambda : (\gamma_1)_k * (\gamma_2)_l \to J_\rho^{-L}, \quad L \gg 0.$$

For $\lambda > 0$ large and $r \in [0,1]$, we define the parameters

$$\lambda_{1,r} = (1-r)\lambda, \quad \lambda_{2,r} = r\lambda. \tag{5.1}$$

We introduce next $\Phi_\lambda(\zeta) = \varphi_{\lambda,\zeta}$ whose components are defined by

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} = \begin{pmatrix} \log \sum_{i=1}^k t_i \Big(\frac{1}{1+\lambda_{1,r}^2 d(x,x_i)^2}\Big)^2 - \frac{1}{2}\log \sum_{j=1}^l s_j \Big(\frac{1}{1+\lambda_{2,r}^2 d(x,y_j)^2}\Big)^2 \\ -\frac{1}{2}\log \sum_{i=1}^k t_i \Big(\frac{1}{1+\lambda_{1,r}^2 d(x,x_i)^2}\Big)^2 + \log \sum_{j=1}^l s_j \Big(\frac{1}{1+\lambda_{2,r}^2 d(x,y_j)^2}\Big)^2 \end{pmatrix}. \tag{5.2}$$

Notice that when $r = 0$ we have that $\lambda_{2,r} = 0$, and therefore, as $\sum_{j=1}^l s_j = 1$, the second terms in both rows are constant, independent of σ_2; a similar consideration holds when $r = 1$. These arguments imply that the function Φ_λ is indeed well defined on $(\gamma_1)_k * (\gamma_2)_l$ (where equivalence relations are used). We have then the following result.

Proposition 5.1 *Suppose that $\rho_1 \in (4k\pi, 4(k+1)\pi)$ and $\rho_2 \in (4l\pi, 4(l+1)\pi)$ and that Σ has positive genus. Then one has*

$$J_\rho(\varphi_{\lambda,\zeta}) \to -\infty \quad \textit{as } \lambda \to +\infty \quad \textit{uniformly in } \zeta \in (\gamma_1)_k * (\gamma_2)_l.$$

Moreover, if $\widetilde{\Psi}$ is as in Proposition 4.2, the composition $\zeta \mapsto \varphi_{\lambda,\zeta} \mapsto \widetilde{\Psi}(\varphi_{\lambda,\zeta})$ is homotopic to the identity on $(\gamma_1)_k * (\gamma_2)_l$ for λ large.

We will not prove this result (referring to [7] for details), but we will limit ourselves to discuss some aspects of the construction and of the estimates. If σ_1 is as above, it turns out that

$$\frac{h_1(x)e^{\varphi_1}}{\int_\Sigma h_1(x)e^{\varphi_1}dV_g} \rightharpoonup \sigma_1 \quad \text{as } \lambda_{1,r} \to +\infty,$$

and similarly for φ_2, replacing h_1 by h_2 and $\lambda_{1,r}$ by $\lambda_{2,r}$. By the way $\widetilde{\Psi}$ is constructed, the latter fact allows to deduce the second statement in Proposition 5.1.

Concerning the estimate of $J_\rho(\varphi_{\lambda,\zeta})$, the most delicate term to understand in (1.3) is the quadratic one in the gradient. Using direct algebraic inequalities, it is possible to prove that

$$|\nabla\varphi_1|(x) \lesssim \min\Big\{\lambda_{1,r}, \frac{4}{d_{1,\min}(x)}\Big\} \quad \text{near } \gamma_1; \quad |\nabla\varphi_1|(x) \lesssim \min\Big\{\lambda_{1,r}, \frac{2}{d_{2,\min}(x)}\Big\} \quad \text{near } \gamma_2,$$

and that, vice versa

$$|\nabla\varphi_2|(x) \lesssim \min\Big\{\lambda_{2,r}, \frac{4}{d_{2,\min}(x)}\Big\} \quad \text{near } \gamma_2, \quad |\nabla\varphi_2|(x) \lesssim \min\Big\{\lambda_{2,r}, \frac{2}{d_{1,\min}(x)}\Big\} \quad \text{near } \gamma_1.$$

In these formulas, we defined $d_{1,\min}(x) = \min_{i=1,\cdots,k} d(x,x_i)$ and $d_{2,\min}(x) = \min_{j=1,\cdots,l} d(x,y_j)$. Moreover it turns out that

$$\nabla\varphi_2 \simeq -\frac{1}{2}\nabla\varphi_1 \quad \text{near } \gamma_1, \quad \nabla\varphi_1 \simeq -\frac{1}{2}\nabla\varphi_2 \quad \text{near } \gamma_2.$$

These two ingredients allow to estimate the desired quantity $\int_\Sigma Q(\varphi_1,\varphi_2)dV_g$. We also notice that, near the peak points x_i and y_j, by the last formula the gradients of φ_1, φ_2 point in opposite directions, and their proportionality turns out to be optimal for keeping Q as small as possible.

We consider now the assumptions of Theorem 1.4, and for simplicity here we limit ourselves to the case $k = 1$, referring to [24] when $k \geq 2$.

When $k = 1$, we wish to parametrize the test functions on the set $(\Sigma * \Sigma) \setminus \widehat{S}$ (see Proposition 4.3). Indeed, as the latter set is not compact, it is convenient to consider, for $\nu > 0$ small, a deformation retract of $(\Sigma * \Sigma) \setminus \widehat{S}$ onto the compact set $\mathcal{X}_\nu$, corresponding to $(\Sigma * \Sigma)$ with a ν-neighbourhood of $\widehat{S}$ removed.

For $\zeta = (\delta_{x_1}, \delta_{x_2}, r) \in \mathcal{X}_\nu$, and $\lambda > 0$ define $\lambda_{1,r}, \lambda_{2,r}$ as in (5.1), and then the test functions $\widehat{\varphi}_{\lambda,\zeta} := (\widehat{\varphi}_1, \widehat{\varphi}_2)$ whose expression is

$$\widehat{\varphi}_1(y) = \log \frac{1 + \lambda_{2,r}^{-2} d(x_2, y)^2}{(1 + \lambda_{1,r} d(x_1, y)^2)^2}, \quad \widehat{\varphi}_2(y) = \log \frac{1 + \lambda_{1,r} d(x_1, y)^2}{(1 + \lambda_{2,r} d(x_2, y)^2)^2}. \tag{5.3}$$

By construction, this map $\widehat{\Phi}_\lambda$ is well defined on $\mathcal{X}_\nu$, and moreover, when one of the parameters t_i is greater than δ these resemble the previous ones.

We have next the counterpart of Proposition 5.1.

Proposition 5.2 *Suppose $\rho_1, \rho_2 \in (4\pi, 8\pi)$. Then one has*

$$J_\rho(\varphi_{\lambda,\zeta}) \to -\infty \quad \text{as } \lambda \to +\infty \quad \text{uniformly in } \zeta \in \mathcal{X}_\nu.$$

Moreover, if $\widehat{\Psi}$ is as in Proposition 4.3, the composition $\zeta \mapsto \varphi_{\lambda,\zeta} \mapsto \widehat{\Psi}(\varphi_{\lambda,\zeta})$ is homotopic to the identity on $\mathcal{X}_\nu$ for λ large.

5.2 The min-max argument

We are now in position to introduce the variational scheme used to prove existence.

Proof of Theorem 1.3 By Proposition 5.1, given any $L > 0$, there exists λ so large that $J_\rho(\varphi_{\lambda,\zeta}) < -L$ for any $\zeta \in (\gamma_1)_k * (\gamma_2)_l$. We choose L so large that Proposition 4.2 applies. We then have that the following composition:

$$(\gamma_1)_k * (\gamma_2)_l \xrightarrow{\Phi_\lambda} J_\rho^{-L} \xrightarrow{\widetilde{\Psi}} (\gamma_1)_k * (\gamma_2)_l$$

is homotopic to the identity map. In this situation, it is said that the set J_ρ^{-L} dominates $(\gamma_1)_k * (\gamma_2)_l$ (see [23, p. 528]). Since $(\gamma_1)_k * (\gamma_2)_l$ is not contractible, this implies that

$$\Phi_\lambda((\gamma_1)_k * (\gamma_2)_l) \text{ is not contractible in } J_\rho^{-L}.$$

Moreover, we can take λ larger so that $\Phi_\lambda((\gamma_1)_k * (\gamma_2)_l) \subset J_\rho^{-2L}$.

Define the topological cone with basis $(\gamma_1)_k * (\gamma_2)_l$ via the equivalence relation

$$\mathcal{C} = \frac{(\gamma_1)_k * (\gamma_2)_l \times [0,1]}{(\gamma_1)_k * (\gamma_2)_l \times \{0\}}.$$

Notice that, since $(\gamma_1)_k * (\gamma_2)_l \simeq S^{2k+2l-1}$, $\mathcal{C}$ is homeomorphic to a Euclidean ball of dimension $2k + 2l$.

We now define the min-max value

$$m = \inf_{\xi \in \Gamma} \max_{u \in \mathcal{C}} J(\xi(u)),$$

where

$$\Gamma = \{\xi : \mathcal{C} \to H^1(\Sigma) \times H^1(\Sigma) : \ \xi(\zeta) = \varphi_{\lambda,\zeta} \ \forall \, \zeta \in \partial\mathcal{C}\}. \tag{5.4}$$

Observe that $t\Phi_\lambda : \mathcal{C} \to H^1(\Sigma) \times H^1(\Sigma)$ belongs to Γ, so this is a non-empty set. Moreover,

$$\sup_{\zeta \in \partial\mathcal{C}} J_\rho(\xi(\zeta)) = \sup_{\zeta \in (\gamma_1)_k * (\gamma_2)_l} J_\rho(\varphi_{\lambda,\zeta}) \leq -2L.$$

We now show that $m \geq -L$. Indeed, $\partial\mathcal{C}$ is contractible in $\mathcal{C}$, and hence in $\xi(\mathcal{C})$ for any $\xi \in \Gamma$. Since $\partial\mathcal{C}$ is not contractible in J_ρ^{-L}, we conclude that $\xi(\mathcal{C})$ is not contained in J_ρ^{-L}. Being this valid for any arbitrary $\xi \in \Gamma$, we conclude that $m \geq -L$. The above argument applies when slightly varying ρ, so we can then apply Proposition 2.1.

Proof of Theorem 1.4 We proceed next similarly to the previous case, restricting ourselves to considering $k = 1$. Let $\overline{\mathcal{X}}_\nu$ denote the topological cone over $\mathcal{X}_\nu$, namely

$$\overline{\mathcal{X}}_\nu = \frac{\mathcal{X}_\nu \times [0,1]}{\mathcal{X}_\nu \times \{1\}}.$$

We choose $L > 0$ so large that Proposition 4.3 applies and then λ so large that, by Proposition 5.2, the supremum of J_ρ on the image of $\widehat{\Phi}_\lambda$ (see the notation after (5.3)) is less than $-2L$.

Consider then the class of maps

$$\Gamma = \{\eta : \overline{\mathcal{X}}_\nu \to H^1(\Sigma) \times H^1(\Sigma) : \eta \text{ is continuous and } \eta(\cdot \times \{0\}) = \varphi_{(\vartheta_1,\vartheta_2)} \text{ on } \mathcal{X}_\nu\}. \tag{5.5}$$

Similarly to the previous case, we have that the set Γ is non-empty and moreover, letting

$$\alpha = \inf_{\eta \in \Gamma} \sup_{m \in \overline{\mathcal{X}}_\nu} J_\rho(\eta(m)),$$

one has

$$\alpha > -\frac{3}{2}L.$$

Indeed, assuming by contradiction that $\alpha \leq -\frac{3}{2}L$, there would be $\eta \in \Gamma$ such that $\sup_{m \in \overline{\mathcal{X}}_\nu} J_\rho(\eta(m)) \leq -\frac{6}{5}L$. Then, letting R_ν denote a retraction of $(\Sigma * \Sigma) \setminus \widehat{S}$ onto $\mathcal{X}_\nu$, writing $m = (\vartheta, s)$ ($\vartheta \in \mathcal{X}_\nu$), the map

$$s \mapsto R_\nu \circ \widehat{\Psi} \circ \eta(\cdot, s)$$

would be a homotopy in $\mathcal{X}_\nu$ between $R_\nu \circ \widehat{\Psi} \circ \varphi_{(\vartheta_1,\vartheta_2)}$ and a constant map.

This fact is indeed impossible since $\mathcal{X}_\nu$ is non-contractible. The proof of this fact is given in the appendix of [38], while here we limit ourselves to describe the case when Σ is a sphere. Indeed in this situation the set $\Sigma * \Sigma \simeq S^2 * S^2$ is homeomorphic to S^5, while $\mathcal{X}_\nu$ is homeomorphic to the product S^5 with a two-dimensional sphere removed. This latter set has a non-vanishing second homology group. Finally, we find that $\alpha > -\frac{3}{2}L$, which is the desired conclusion.

As before, the reasoning applies when slightly varying ρ, so we can then apply Proposition 2.1.

References

[1] Aubin, T., Some Nonlinear Problems in Differential Geometry, Springer-Verlag, New York, 1998.

[2] Bahri, T. and Coron, J. M., On a nonlinear elliptic equation involving the critical Sobolev exponent: The effect of the topology of the domain, *Comm. Pure Appl. Math.*, **41**, 1988, 253–294.

[3] Bartolucci, D., De Marchis, F. and Malchiodi, A., Supercritical conformal metrics on surfaces with conical singularities, *Int. Math. Res. Not.*, **2011**(24), 2011, 5625–5643.

[4] Bartolucci, D. and Malchiodi, A., An improved geometric inequality via vanishing moments, with applications to singular Liouville equations, *Comm. Math. Phys.*, **322**(2), 2013, 415–452.

[5] Battaglia, L., Existence and multiplicity result for the singular Toda system, *J. Math. Anal. Appl.*, **424**(1), 2015, 49–85.

[6] Battaglia, L. and Mancini, G., A note on compactness properties of the singular Toda system, *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, **26**(3), 2015, 299–307.

[7] Battaglia, L., Jevnikar, A., Malchiodi, A. and Ruiz, D., A general existence result for the Toda system on compact surfaces, *Adv. in Math.*, **285**, 2015, 937–979.

[8] Bolton, J. and Woodward, L. M., Some geometrical aspects of the 2-dimensional Toda equations (in Geometry), Topology and Physics, Campinas, 1996, 69–81; de Gruyter, Berlin, 1997.

[9] Brezis, H. and Merle, F., Uniform estimates and blow-up behavior for solutions of $-\Delta u = V(x)\mathrm{e}^u$ in two dimensions, *Commun. Partial Differ. Equations*, **16**(8–9), 1991, 1223–1253.

[10] Calabi, E., Isometric imbedding of complex manifolds, *Ann. Math.*, **58**(2), 1953, 1–23.

[11] Carlotto, A. and Malchiodi, A., Weighted barycentric sets and singular Liouville equations on compact surfaces, *J. Funct. Anal.*, **262**(2), 2012, 409–450.

[12] Chanillo, S. and Kiessling, M., Rotational symmetry of solutions of some nonlinear problems in statistical mechanics and in geometry, *Comm. Math. Phys.*, **160**, 1994, 217–238.

[13] Chen, W. X. and Li, C., Prescribing Gaussian curvatures on surfaces with conical singularities, *J. Geom. Anal.*, **1**(4), 1991, 359–372.

[14] Chen, C. C. and Lin, C. S., Topological degree for a mean field equation on Riemann surfaces, *Comm. Pure Appl. Math.*, **56**(12), 2003, 1667–1727.

[15] Chen, X. X., Remarks on the existence of branch bubbles on the blowup analysis of equation $-\Delta u = \mathrm{e}^{2u}$ in dimension two, *Commun. Anal. Geom.*, **7**(2), 1999, 295–302.

[16] Chern, S. S. and Wolfson, J. G., Harmonic maps of the two-sphere into a complex Grassmann manifold, II, *Ann. of Math.*, **125**(2), 1987, 301–335.

[17] Daprile, T., Pistoia, A. and Ruiz, D., Asymmetric blow-up for the $SU(3)$ Toda system, *J. Funct. Anal.*, **271**(3), 2016, 495–531.

[18] Ding, W., Jost, J., Li, J. and Wang, G., The differential equation $\Delta u = 8\pi - 8\pi h \mathrm{e}^u$ on a compact Riemann surface, *Asian J. Math.*, **1**, 1997, 230–248.

[19] Ding, W., Jost, J., Li, J. and Wang, G., Existence results for mean field equations, *Ann. Inst. Henri Poincaré, Anal. Non Linèaire*, **16**(5), 1999, 653–666.

[20] Djadli, Z., Existence result for the mean field problem on Riemann surfaces of all genus, *Comm. Contemp. Math.*, **10**(2), 2008, 205–220.

[21] Djadli, Z. and Malchiodi, A., Existence of conformal metrics with constant Q-curvature, *Ann. Math.*, **168**(3), 2008, 813–858.

[22] Dunne, G., Self-dual Chern-Simons Theories, Lecture Notes in Physics, Vol. 36, Springer-Verlag, Berlin, 1995.

[23] Hatcher, A., Algebraic Topology, Cambridge University Press, Cambridge, 2002.

[24] Jevnikar, A., Kallel, S. and Malchiodi, A., A topological join construction and the Toda system on compact surfaces of arbitrary genus, *Anal. PDE*, **8**(8), 2015, 1963–2027.

[25] Jost, J., Lin, C. S. and Wang, G., Analytic aspects of the Toda system II, Bubbling behavior and existence of solutions, *Comm. Pure Appl. Math.*, **59**, 2006, 526–558.

[26] Jost, J. and Wang, G., Analytic aspects of the Toda system I, A Moser-Trudinger inequality, *Comm. Pure Appl. Math.,* **54**, 2001, 1289–1319.

[27] Kallel, S. and Karoui, R., Symmetric joins and weighted barycenters, *Advanced Nonlinear Studies*, **11**, 2011, 117–143.

[28] Li, J. and Li, Y., Solutions for Toda systems on Riemann surfaces, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5), **4**(4), 2005, 703–728.

[29] Li, Y. Y., Harnack type inequality: The method of moving planes, *Commun. Math. Phys.*, **200**(2), 1999, 421–444.

[30] Li, Y. Y. and Shafrir, I., Blow-up analysis for solutions of $-\Delta u = V\mathrm{e}^u$ in dimension two, *Indiana Univ. Math. J.*, **43**(4), 1994, 1255–1270.

[31] Lin, C. S., Wei, J. and Yang, W., Degree counting and shadow system for $SU(3)$ Toda system: One bubbling, 2014, preprint. https://arxiv.org/abs/1408.5802

[32] Lin, C. S., Wei, J. and Zhao, C., Sharp estimates for fully bubbling solutions of a $SU(3)$ Toda system, *Geom. Funct. Anal.*, **22**(6), 2012, 1591–1635.

[33] Lin, C. S. and Zhang, L., A topological degree counting for some Liouville systems of mean field type, *Comm. Pure Appl. Math.*, **64**, 2011, 556–590.

[34] Lucia, M., A deformation lemma with an application to a mean field equation, *Topol. Methods Nonlinear Anal.*, **30**(1), 2007, 113–138.

[35] Malchiodi, A., Morse theory and a scalar field equation on compact surfaces, *Adv. Diff. Eq.*, **13**, 2008, 1109–1129.

[36] Malchiodi, A. and Ndiaye, C. B., Some existence results for the Toda system on closed surfaces, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.,* **18**(4), 2007, 391–412.

[37] Malchiodi, A. and Ruiz, D., New improved Moser-Trudinger inequalities and singular Liouville equations on compact surfaces, *Geom. Funct. Anal.*, **21**(5), 2011, 1196–1217.

[38] Malchiodi, A. and Ruiz, D., A variational analysis of the Toda system on compact surfaces, *Comm. Pure Appl. Math.*, **66**(3), 2013, 332–371.

[39] Malchiodi, A. and Ruiz, D., On the Leray-Schauder degree of the Toda system on compact surfaces, *Proc. Amer. Math. Soc.*, **143**(7), 2015, 2985–2990.

[40] Nolasco, M. and Tarantello, G., On a sharp Sobolev-type inequality on two-dimensional compact manifolds, *Arch. Ration. Mech. Anal.*, **145**, 1998, 161–195.

[41] Prajapat, J. and Tarantello, G., On a class of elliptic problems in R^2: Symmetry and uniqueness results, *Proc. Roy. Soc. Edinburgh, Ser. A*, **131**, 2001, 967–985.

[42] Struwe, M., On the evolution of harmonic mappings of Riemannian surfaces, *Comment. Math. Helv.*, **60**, 1985, 558–581.

[43] Struwe, M. and Tarantello, G., On multivortex solutions in Chern-Simons gauge theory, *Boll. Unione Mat. Ital., Sez. B, Artic. Ric. Mat.*, **8**(1), 1998, 109–121.

[44] Tarantello, G., Self-Dual Gauge Field Vortices: An Analytical Approach, PNLDE 72, Birkhäuser Boston, Boston, MA, 2007.

[45] Yang, Y., Solitons in Field Theory and Nonlinear Analysis, Springer-Verlag, New York, 2001.