

# Low-dimensional quadratic relative character varieties for the projective line minus three points

Carlos Simpson

## ▶ To cite this version:

Carlos Simpson. Low-dimensional quadratic relative character varieties for the projective line minus three points. Science China Mathematics, In press, 10.1007/s11425-024-2330-8. hal-04705055

## HAL Id: hal-04705055 https://hal.science/hal-04705055v1

Submitted on 22 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## LOW-DIMENSIONAL QUADRATIC RELATIVE CHARACTER VARIETIES FOR THE PROJECTIVE LINE MINUS THREE POINTS

## CARLOS SIMPSON

Dedicated to my friend Gang Xiao

ABSTRACT. We show that relative character varieties having a fixed dimension d>0, for irreducible orthogonal or symplectic local systems of rank r on  $\mathbb{P}^1-\{0,1,\infty\}$ , are isomorphic via Katz's middle convolution to ones with a bound on the rank  $r\leq R(d)$ .

#### 1. Introduction

In this paper we consider the quadratic relative character varieties of local systems over  $\mathbb{P}^1 - \{0, 1, \infty\}$ . A quadratic local system is an irreducible local system  $\mathcal{F}$  such that  $\mathcal{F} \cong \mathcal{F}^{\vee}$ . In this case the bilinear form  $(\ ,\ ): \mathcal{F} \otimes \mathcal{F} \to \underline{\mathbb{C}}$  is unique up to scalars, nondegenerate, and it is either symmetric or antisymmetric.

Let r denote the rank of  $\mathcal{F}$  and introduce the sign  $\varepsilon = \pm 1$  so that the form is  $\varepsilon$ -symmetric, i.e. if  $\varepsilon = 1$  it is symmetric and if  $\varepsilon = -1$  it is antisymmetric.

The monodromy representation of  $\mathcal{F}$  is

$$\rho: \pi_1(\mathbb{P}^1 - \{0, 1, \infty\}, x) \to G_{\varepsilon}$$

where x denotes a basepoint and  $G_{\varepsilon} = O(r)$  (if  $\varepsilon = 1$ ) or  $G_{\varepsilon} = Sp(r)$  (if  $\varepsilon = -1$ ). Let  $C_1, C_2, C_3$  denote the conjugacy classes of the monodromy transformations around the singular points  $t_1, t_2, t_3 = 0, 1, \infty$ .

The quadratic relative character variety  $\mathscr{X}^{\text{irred}}(r, \varepsilon, C_1, C_2, C_3)$  of irreducible local systems is defined to be the algebraic space quotient of  $\mathbf{Irrep}(r, \varepsilon, C_1, C_2, C_3)$  by the action of  $G_{\varepsilon}$ , where

$$\mathbf{Irrep}(r, \varepsilon, C_1, C_2, C_3) \subset C_1 \times C_2 \times C_3$$

is the variety of triples  $(A_1, A_2, A_3)$  of matrices in the given conjugacy classes in  $G_{\varepsilon}$  with  $A_1A_2A_3=1$  and that generate irreducible representations of the

 $<sup>2020\</sup> Mathematics\ Subject\ Classification.$  Primary 14HD20; Secondary 14H30, 14L35, 20G20.

Key words and phrases. fundamental group, representation, middle convolution, relative character variety, parabolic Higgs bundle, orthogonal group, symplectic group, conjugacy class.

To appear in Science China Mathematics. The original publication will be available at www.scichina.com and www.springerlink.com.

fundamental group in GL(r). For the variety of irreducible points we have  $\dim \mathscr{X}^{irred}(r, \varepsilon, C_1, C_2, C_3) = \dim(C_1) + \dim(C_2) + \dim(C_3) - 2\dim(G_{\varepsilon})$ .

Katz defined an operation of middle convolution on  $GL(r, \mathbb{Q}_l)$ -local systems, that was then described algebraically for  $GL(r, \mathbb{C})$ -local systems by Dettweiler-Reiter [23]. We'll show here that certain compositions of two middle convolutions may be interpreted as middle convolution with rank 2 local systems that preserve quadraticity. This gives isomorphisms between the quadratic relative character varieties for different ranks and conjugacy classes. As the conjugacy classes in  $G_{\varepsilon}$  are determined by their corresponding classes in GL(r), Katz's formulas for the variation of conjugacy classes under middle convolution yield formulas for these operations here.

**Theorem 1.1.** There are a series of inequalities given in Lemmas 4.1, 4.2, 4.3, 4.4, 4.5 below, such that any quadratic local system can be brought by quadratic middle convolution operations to one that satisfies these inequalities, a condition we call numerically MC-minimal. For a given dimension  $d \geq 2$  there is a bound R(d) such that any numerically MC-minimal irreducible quadratic local system has rank  $r \leq R(d)$ . In particular, for a given dimension  $d \geq 2$  there are up to isomorphism only finitely many families—the parameters being the choices of eigenvalues—of quadratic relative character varieties of that dimension.

This property is relatively easy to see for the group GL(r) using an easier version of what we will be doing here, see the discussion in [58, Section 2]. The main point is that the dimension is equal to the area left over after removing squares whose sizes are the parts in the partitions, from a big square of size  $r \times r$ . After reduction using middle convolution the squares fit into vertical rectangles (what we will call the non-overlapping case below). As r becomes large, a bound saying that the leftover area is less than a fixed constant d constrains the configuration. This is similar to our discussion in Sections 5 and 6 below, and the reader is invited to do it as an exercise.

The difficulty to be treated in the present paper stems from the fact that the dimensions of the orthogonal or symplectic groups G contain a linear term, indeed

$$\dim(O(r)) = \frac{r^2 - r}{2}, \quad \dim(Sp(r)) = \frac{r^2 + r}{2}.$$

There are similar linear corrections to the dimensions of the conjugacy classes  $C_i$ . The leftover area after subtracting the smaller boxes from the  $r \times r$  box therefore satisfies an estimate of the form  $\leq r + 2d$ . A more intricate argument becomes necessary.

A similar theorem has recently been proven by Domokos for quiver representations [27].

Before getting there, we review basic facts about middle convolution, the parametrisation and dimensions of conjugacy classes in the orthogonal or symplectic groups, and prove the lemmas showing that reductions using quadratic middle convolution operations can be used to attain the bounds of numerical MC-minimality.

1.1. **Motivation.** My motivation for this project was the idea of looking for natural constructions of algebraic surfaces within the theory of character varieties and parabolic Higgs bundles, with a view towards gaining examples with automorphisms.

Suppose that the monodromy transformations are semisimple with eigenvalues in the unit circle  $U(1) \subset \mathbb{C}^*$ . This is compatible with the quadratic condition. The paired eigenvalues have nontrivial imaginary part and come in complex conjugate pairs, and the half-unital eigenvalues  $\pm 1$  are the real eigenvalues. The relative character variety  $\mathscr{X}^{\text{irred}}(r, \varepsilon, C_1, C_2, C_3)$  has a Hitchin-type  $Dolbeault\ version$  defined as the moduli space  $M_{Dol}^{\text{stab}}(r, \varepsilon, C_1, C_2, C_3)$  of stable strictly parabolic Higgs bundles such that the dimensions of the graded pieces in the parabolic flags are the parts of the partitions associated to  $C_i$ , that are self-dual with type  $\varepsilon$ .

We can only mention a selection of references. For moduli spaces of parabolic Higgs or vector bundles these include [1, 15, 40, 50, 51, 52, 55, 68]. For the theory of G-bundles and parabolic G-bundles these include [3, 9, 10, 11, 13]. For G-local systems and connections, see for example [16, 19, 36, 57, 60]. Parahoric structures, generalizing parabolic structures for other groups such as  $G_{\varepsilon}$ , are studied in [5, 7, 14, 44, 54]. The theory of spectral varieties [37] becomes more subtle in the cases of G-bundles and parabolic structures [4, 34, 35, 56, 64]. A few references for the Kobayashi-Hitchin correspondence [38] in this context would include [2, 18, 21, 53]. Although we only consider complex groups in the present paper, the vast topic of real structures [12, 17, 31] is closely related.

The relative character variety is homeomorphic to  $M_{Dol}^{\mathrm{stab}}(r,\varepsilon,C_1,C_2,C_3)$ . In turn  $M_{Dol}^{\mathrm{stab}}(r,\varepsilon,C_1,C_2,C_3)$  is an open subset of the moduli space  $M_{Dol}(r,\varepsilon,C_1,C_2,C_3)$  of S-equivalence classes of semistable objects. If we assume that the eigenvalues are roots of unity then both the relative character variety and the Hitchin moduli space may be viewed as components of the moduli spaces of local systems or Higgs bundles on a Deligne-Mumford root stack over  $\mathbb{P}^1$  with three stacky points. The theory in the compact case then applies [59].

This bigger space has a  $\mathbb{C}^*$ -action. The limits of orbits as  $t \to 0$  exist, and for starting points in a dense Zariski open subset, the limits are contained in the component of the fixed-point set of highest dimension, one-half the global dimension. We call this the *bottom component* since it is the minimum locus of Hitchin's energy functional.

Over smooth compact curves, the bottom component is the moduli space of unitary local systems. Over orbifolds, or more generally in the parabolic case, the bottom component can sometimes be a moduli space of unitary local systems, or within the Hitchin space it is the moduli space of semistable parabolic bundles. The conditions for existence of a stable parabolic bundle in the bottom component are complicated (see [8] and the references there).

But the bottom component can also be a moduli space of variations of Hodge structure of some other type. Examples giving compact components of real group character varieties were constructed by Tholozan and Toulisse [62].

Natural automorphisms of the relative character variety will typically be "motivic" in origin, and will consequently respect the full nonabelian Hodge package. In particular, they give automorphisms of  $M_{Dol}(r, \varepsilon, C_1, C_2, C_3)$  respecting the  $\mathbb{C}^*$ -action, and therefore induce automorphisms of the bottom component. In higher genus Alfaya-Gómez show a converse [1].

These moduli spaces have hyperkähler structures [38, 29, 30, 39, 52]. In particular, the complex dimensions of the relative character varieties are even, thus the smallest positive dimensions are  $2, 4, \ldots$  The bottom components then have dimensions  $1, 2, \ldots$ 

In conclusion, for the case when  $\mathscr{X}^{\text{irred}}(r,\varepsilon,C_1,C_2,C_3)$  has dimension d=4 we are going to get surfaces, the bottom components, that can have groups of natural automorphisms.

Our motivation was to begin to go towards the classification of such examples. From the finiteness of Theorem 1.1, we should be able to compute the list of possibilities with dimension d=4 (see Problem 8.2). The question of non-emptiness of the moduli spaces is not treated here, but one may expect that techniques such as cyclic Higgs bundles [21, 58] the very good property [60], quiver techniques [2, 20, 27, 65], Kostov's explicit techniques [42], the results of Pandey [54], Wen [65] and others, should allow to treat the existence question to a large extent.

A next problem will be to obtain precise descriptions of the bottom components as explicit moduli spaces of quadratic parabolic vector bundles or Hodge bundles.

Automorphisms can come from several different places, including automorphisms of the base curve—in this case  $\mathbb{P}^1 - \{0, 1, \infty\}$ , but also automorphisms coming from tensoring with rank 1 local systems, and automorphisms coming from (compositions of) quadratic middle convolution operations, in the case that those preserve the dimensions and conjugacy classes.

Although the moduli spaces do not often look to be of general type, we note that the automorphisms will also preserve the *wobbly loci* [28, 46] and the pair consisting of the bottom component together with its wobbly locus is likely going to be of general type.

It will be interesting to see whether there are examples that come close to Xiao's bound [66] [67] for the number of automorphisms as a function of the numerical properties.

#### 2. Quadratic local systems

We consider local systems on  $C = \mathbb{P}^1 - \{t_1, \dots, t_k\}$ . For the main theorem k = 3 but the preliminary remarks can take place for any k.

Following Pandey [54, Definition 2.0.2], say that an irreducible local system is *quadratic* if it is isomorphic to its dual.

Suppose  $\mathcal{F}$  is irreducible and quadratic. Then  $Hom(\mathcal{F}, \mathcal{F}^{\vee}) \cong End(\mathcal{F}) \cong \mathbb{C}$  so there is a unique bilinear pairing

$$(\ ,\ ): \mathcal{F} \times \mathcal{F} \to \underline{\mathbb{C}}.$$

The pairing (u, v)' := (v, u) is therefore the same, up to a scalar, in other words there is  $\varepsilon \in \mathbb{C}^*$  with

$$(v, u) = \varepsilon(u, v).$$

As  $(u,v) = \varepsilon(v,u) = \varepsilon^2(u,v)$  we have  $\varepsilon^2 = 1$  so  $\varepsilon = \pm 1$ . Thus if  $\varepsilon = 1$  the pairing is a nondegenerate symmetric bilinear form and if  $\varepsilon = -1$  it is a nondegenerate antisymmetric bilinear form. Fixing a standard basis over the basepoint  $x \in C$ , the monodromy representation

$$\rho: \pi_1(C,x) \to GL(r,\mathbb{C})$$

factors through the group  $G_{\varepsilon} \subset GL(r,\mathbb{C})$  with  $G_{\varepsilon} = O(r,\mathbb{C})$  if  $\varepsilon = 1$  and  $G_{\varepsilon} = Sp(r,\mathbb{C})$  if  $\varepsilon = -1$ . In the latter case r has to be even.

If we start with a quadratic local system and do a middle convolution with a quadratic convolution kernel L that is MC-invertible (i.e. convolution with L admits an inverse), then the result is again a quadratic local system (Proposition 3.2).

**Definition 2.1.** We say that an irreducible quadratic local system is MC-minimal if the rank does not decrease under middle convolution with quadratic MC-invertible convolution kernels.

We are going to look for restrictions on the monodromy data of MC-minimal local systems, in the case k=3. Considering various kinds of middle convolutions with local systems L will give a collection of inequalities on the eigenvalue multiplicities. Middle convolution will be discussed in more detail in Section 3.

2.1. The quadratic relative character varieties. Suppose  $C_i^{GL}$  are conjugacy classes in the general linear group GL(r). Let  $\overline{C}_i^{GL}$  denote their closures. Let  $\mathscr{X}(r,\overline{C}_1^{GL},\ldots,\overline{C}_k^{GL})$  denote the closed relative character variety, the closed subset of the character variety of the free group  $\pi_1(C,x)$  consisting of representations whose monodromy transformations at the points  $t_i$  are in  $\overline{C}_i^{GL}$ . This is an affine variety, defined as the GIT quotient of the closed affine subvariety of  $\overline{C}_1^{GL} \times \cdots \times \overline{C}_k^{GL}$ , consisting of tuples whose product is 1, by the action of GL(r). Recall from the theory of the character variety [47] that there is a Zariski open subset whose points parametrize isomorphism classes of irreducible representations. In this subset we can further take a

Zariski open subset whose points parametrize representations such that the monodromy transformations are in  $C_i$ . This yields a Zariski open subset

$$\mathscr{X}^{\mathrm{irred}}(r, C_1^{GL}, \dots, C_k^{GL}) \subset \mathscr{X}(r, \overline{C}_1^{GL}, \dots, \overline{C}_k^{GL})$$

whose points parametrize isomorphism classes of irreducible local systems on  $C = \mathbb{P}^1 - \{t_1, \dots, t_k\}$  whose local monodromy transformations are in the conjugacy classes  $C_i$ .

Suppose now given  $\varepsilon = \pm 1$  and suppose given conjugacy clases  $C_1, \ldots, C_k$  in the group  $G_{\varepsilon} \subset GL(r)$  of automorphisms preserving the standard  $\varepsilon$ -symmetric bilinear form. Let  $C_i^{GL}$  denote the conjugacy classes they generate in GL(r). These have the property that they are self-dual. Recall [43, 63, 32] that the specification of  $C_i^{GL}$  and  $\varepsilon$  determines uniquely the conjugacy class  $C_i \subset G_{\varepsilon}$ .

**Remark 2.2.** Note that there are conditions on the  $C_i^{GL}$ , depending on  $\varepsilon$ , for the existence of  $C_i$  [43, Section 2.2], [63, Theorem 1.3.4] [61]. In particular,  $C_i$  might exist for one value of  $\varepsilon$  but not for the other one.

**Proposition 2.3.** There is a quasiprojective variety whose points parametrize  $G_{\varepsilon}$ -isomorphism classes of representations

$$\mathscr{X}^{\mathrm{irred}}(r, \varepsilon, C_1, \dots, C_k) := \frac{\ker \left( (C_1 \times \dots \times C_k)^{\mathrm{irred}} \to G_{\varepsilon} \right)}{G_{\varepsilon}}$$

where the superscript irred denotes representations that are irreducible in GL(r). The map

$$\mathscr{X}^{\mathrm{irred}}(r, \varepsilon, C_1, \dots, C_k) \to \mathscr{X}^{\mathrm{irred}}(r, C_1^{GL}, \dots, C_k^{GL})$$

is a closed embedding. The union of the (one or two cf Remark 2.2) images is the fixed point set of the involution  $\mathcal{F} \mapsto \mathcal{F}^{\vee}$  on the GL(r) character variety.

*Proof.* The construction is classical in the same way as was described above for GL(r). The points of the GIT quotient parametrize closed  $G_{\varepsilon}$ -orbits, and a  $G_{\varepsilon}$ -local system whose associated GL(r)-local system is irreducible, is in a closed  $G_{\varepsilon}$ -orbit (the following argument may be used to show this).

Suppose given two representations  $\rho, \rho': \pi_1(C, x) \to G_{\varepsilon}$  which are conjugate as irreducible representations into GL(r). Thus, there is  $g \in GL(r)$  such that  $g\rho'(a) = \rho(a)g$  for all  $a \in \pi_1(C, x)$ . Let  $g^*$  denote the adjoint with respect to the bilinear form. Then for any vectors u, v

$$(u, \rho'(a)^{-1}g^*g\rho'(a)v) = (g\rho'(a)u, g\rho'(a)v) = (\rho(a)gu, \rho(a)gv) = (u, g^*gv)$$

so  $\rho'(a)^{-1}g^*g\rho'(a) = g^*g$ . This holds for any a and since  $\rho'$  is irreducible it implies that  $g^*g$  is a scalar  $\lambda \mathbf{1}$ . Let  $\alpha$  be a square-root of  $\lambda^{-1}$ , then  $(\alpha g)^* = (\alpha g)$ . Thus  $\alpha g \in G_{\epsilon}$  intertwines  $\rho$  and  $\rho'$ .

This shows that the map from the  $G_{\varepsilon}$  to the GL(r) character variety is an inclusion on the level of points. We also know that it is a finite map

in general (the Zariski open subset for  $G_{\varepsilon}$  is pulled back from the one for GL(r)), and both varieties are smooth, so it is a closed embedding.

The image is contained in the fixed point set of  $\mathcal{F} \mapsto \mathcal{F}^{\vee}$ , and on the other hand any point in the fixed point set is by definition a quadratic local system of one of the two types  $\varepsilon = 1$  or  $\varepsilon = -1$ . This gives the last statement. Notice that there might be cases where the subset is empty for one of the  $\varepsilon$  but not for the other one. Also, we are not making any statement here about connectedness of these pieces of the fixed point set.

**Lemma 2.4.** The dimension the quadratic relative character variety, if it is nonempty, is

$$\dim \mathscr{X}^{\mathrm{irred}}(r, \varepsilon, C_1, \dots, C_k) = \dim(C_1) + \dots + \dim(C_k) - 2\dim(G_{\varepsilon}).$$

*Proof.* We give a conceptual proof that doesn't depend on calculations with constructible sheaves, however one may also refer to [41]. Irreducibility of the local systems will imply the vanishing of the obstruction space; note that without this assumption it is not easy to get a bound on the size of the obstruction space, see [33].

Imagine the same situation on a curve C whose compactification  $\overline{C}$  has high genus g but still with k punctures. The formula for the dimension is the same, plus  $2g\dim(G_{\varepsilon})$ . Also, the formula for the dimension is  $h^1$  of a certain sheaf controlling deformations with fixed conjugacy classes, whose local structure at the singularities is the same for curves of any genus. The obstruction space is  $h^2$  of this sheaf. There is a Verdier duality between  $h^2$ and  $h^0$ . The  $h^0$  in question is the dimension of the stabilizer of the representation in  $G_{\varepsilon}$ , as can be noted from the fact that both the dimension of the stabilizer and the obstruction space are 0 when  $q \gg 0$ , plus the fact that whatever the relationship is between these two, it comes from local contributions at the punctures so it is independent of g. We conclude that the obstruction space vanishes for quadratic local systems over  $\mathbb{P}^1 - \{t_1, \dots, t_k\}$ that are irreducible and hence have 0-dimensional stabilizer group. Then by the same type of argument, the  $h^1$  is equal to the Euler characteristic of this sheaf which again decomposes as  $2q\dim(G_{\varepsilon})$  plus a term that depends only on the local structure. This is therefore equal to the stated formula when q=0.

2.2. Conjugacy classes in the orthogonal and symplectic groups. In order to apply Lemma 2.4 we need to know the dimensions of the conjugacy classes. This is classical [61], we follow the references Topley [63] and Kraft-Procesi [43], and try to adopt a notation close to theirs.

Fix  $\varepsilon = \pm 1$ . Let V be an r-dimensional vector space with a nondegenerate  $\varepsilon$ -symmetric bilinear form [43, Section 1]. The group  $G_{\varepsilon}$  of automorphisms of V preserving the form is the orthogonal group if  $\varepsilon = 1$  and the symplectic group if  $\varepsilon = -1$ . We have

$$\dim(G_{\epsilon}) = \frac{1}{2} (r^2 - \varepsilon r).$$

The group  $G_{\epsilon}$  was denoted G(V) in [43] and K in [63].

A conjugacy class of  $G_{\epsilon}$  is determined by the conjugacy class it generates in GL(V) [43, Section 2.2]. This in turn is determined by a partition  $\lambda$  of r and a collection of eigenvalues one for each part of the partition. Nontrivial unipotent parts arise when there are several parts for the same eigenvalue.

The collection of eigenvalues with multiplicities is invariant under the automorphism  $z\mapsto z^{-1}$ . It is convenient to identify two types of eigenvalues that we'll call half-unital, preserved by inversion, and paired, when non-preserved by inversion. The half-unital eigenvalues are either 1 or -1, whereas any other eigenvalue  $a\neq \pm 1$  is paired. When a is paired, the Jordan normal forms at a and  $a^{-1}$  are identified. As the partitions we use are the duals of the Jordan normal form partitions, the pieces of the partitions at a and  $a^{-1}$  are identified.

In our discussion, the phenomenon of nontrivial unipotent parts at paired eigenvalues can safely be ignored, because there will be no difference between parts having the same paired eigenvalues and those having distinct ones in the formulas for the dimension.

There can be two half-unital eigenvalues 1 and -1. We need to consider the case where there are nontrivial unipotent parts at these eigenvalues, even if we originally start with semisimple conjugacy classes, because such unipotent parts can arise during the middle convolution reductions. We note also that recent work in mathematical physics has a focus on nontrivial unipotent parts [6, 64].

We are therefore interested in the dimension of the conjugacy class of a unipotent element of  $G_{\epsilon}$ . For the -1 eigenvalue the dimension formula is the same.

For this section let us now consider a unipotent element with partition  $r = \lambda_1 + \ldots + \lambda_d$  with  $\lambda_1 \geq \ldots \geq \lambda_d$ . The partition  $\eta$  giving the dimensions of the Jordan blocks is the dual of  $\lambda$ , so  $\lambda_j$  is the number of parts of  $\eta$ , i.e. the number of Jordan block sizes, that are  $\geq j$ . In Kraft-Procesi [43, Section 2] the notation is  $\lambda = \hat{\eta}$  and  $|\eta| = r$ .

The formula for twice<sup>1</sup> the dimension of the centralizer Z of an element in the conjugacy class as given in [43, Section 2.4] is

$$2\dim(Z) = \sum_{i} (\hat{\eta}_i)^2 - \varepsilon \# \{j, \, \eta_j \text{ odd} \}.$$

Below we will replace  $\hat{\eta}_i$  by  $\lambda_i$ .

We can count the number of j such that  $\eta_j$  is odd as follows. The number with  $\eta_j = 1$  is  $\lambda_1 - \lambda_2$ . The number with  $\eta_j = 3$  is  $\lambda_3 - \lambda_4$ , and so forth. Thus, the number of indices with  $\eta_j$  odd is

$$(\lambda_1 - \lambda_2) + (\lambda_3 - \lambda_4) + \ldots = -\sum_{i=1}^{n} (-1)^i \lambda_i.$$

<sup>&</sup>lt;sup>1</sup>Dimensions will be multiplied by 2 in order to avoid too many fractions in the notations.

This gives

$$2\dim(Z) = \sum_{i} (\lambda_i)^2 + \varepsilon \sum_{i} (-1)^i \lambda_i.$$

For the case of a conjugacy class  $C_i$  that is not necessarily unipotent, we introduce the following notations. Let  $m_i^j$  be the parts corresponding to paired eigenvalues. They occur in equal pairs consisting of the eigenvalues and their inverses. We count each pair only once, and order them so that  $m_i^0$  is the biggest one.

Let  $e_i^j$  and  $f_i^j$  be the parts of the partitions corresponding to the halfunital eigenvalues  $\sigma_i$  and  $-\sigma_i$  respectively for  $\sigma_i = \pm 1$  (see below). They start with j = 0 and are ordered by

$$e_i^0 \ge e_i^1 \ge \dots, \quad f_i^0 \ge f_i^1 \ge \dots,$$

and choose  $\sigma_i$  to be the eigenvalue such that  $e_i^0 \geq f_i^0$ .

In the semisimple case there are just one multiplicity  $e_i^0$  and  $f_i^0$  for each eigenvalue, and  $\sigma_i$  is chosen as the eigenvalue with the biggest multiplicity  $e_i^0$ .

The partition relation is

$$r = \sum_{j} e_i^j + \sum_{j} f_i^j + 2 \sum_{j} m_i^j$$

where the factor of 2 is because we use  $m_i^j$  for the multiplicities of both paired eigenvalues  $a_i^j$  and  $(a_i^j)^{-1}$ .

We have seen above the formula for the dimension of the centralizer of each unipotent block. These add together to give the dimension of the centralizer  $Z_i$ ,

$$2\dim(Z_i) = 2\sum_{j} (m_i^j)^2 + \sum_{j} (e_i^j)^2 - \varepsilon \sum_{j} (-1)^j e_i^j + \sum_{j} (f_i^j)^2 - \varepsilon \sum_{j} (-1)^j f_i^j.$$

The change of sign for the linear term here is due to the fact that we changed indexing so that it starts with j = 0.

**Proposition 2.5.** Suppose  $\mathcal{F}$  is an irreducible quadratic local system of rank r and of type  $\varepsilon$ . Let k denote the number of singular points. The dimension at  $\mathcal{F}$  of the relative character variety  $\mathscr{X}^{irred}(r, \varepsilon, C_1, \ldots, C_k)$  of type  $\varepsilon$  quadratic local systems with the same local monodromy transformations as  $\mathcal{F}$ , multiplied by 2, is

$$(k-2)r^2 - 2\sum_{i,j}(m_i^j)^2 - \sum_{i,j}(e_i^j)^2 - \sum_{i,j}(f_i^j)^2 - \varepsilon \left( (k-2)r - \sum_{i,j}(-1)^j e_i^j - \sum_{i,j}(-1)^j e_i^j \right).$$

*Proof.* This is Lemma 2.4 rewritten using the expressions for the dimensions of the conjugacy classes discussed above.  $\Box$ 

Note that there is a case where the O(r)-conjugacy class is disconnected and splits into two SO(r)-conjugacy classes, see [43, Section 2.3]. As we are

concerned for the moment with calculating dimensions and getting bounds, we do not consider this distinction here and look at the O(r)-conjugacy class that is determined by the GL(r)-conjugacy class it generates. The disconnectedness phenomenon will of course be interesting when it comes to computing examples.

### 3. MIDDLE CONVOLUTION

In this section we consider the case of local systems over  $C := \mathbb{P}^1 - \{t_1, \ldots, t_k\}$ , whereas in the remainder of the paper k = 3.

Let  $\overline{C} := \mathbb{P}^1$  with its divisor  $D = t_1 + \dots t_k$  of marked points.

Let  $\overline{X}$  be the blow-up of  $\overline{C} \times \overline{C}$  at the points  $(t_i, t_i)$  along the diagonal. Let  $\Delta \subset \overline{X}$  be the divisor consisting of

$$\Delta = T + (H_1 + \ldots + H_k) + (U_1 + \ldots + U_k) + (V_1 + \ldots + V_k)$$

where  $H = H_1 + \ldots + H_k$  is the strict transform of  $D \times \overline{C}$ , V is the strict transform of  $\overline{C} \times D$ , T is the strict transform of the diagonal, and U is the exceptional divisor.

Let  $p, q : \overline{X} \to \overline{C}$  come from the first and second projections respectively. Thus, H is horizontal for q and V is vertical for q.

Let  $j: X \hookrightarrow \overline{X}$  be the complement of  $\Delta$ . Let  $p^o$  and  $q^o$  be the restrictions of p and q to X. Let  $j^o$  denote the restriction of j to the open subset  $\overline{C} \times C \subset X$ , and let q' denote the restriction of q to this open subset. Suppose L is a local system on X that will be called the *convolution kernel*. Define

$$MC(\mathcal{F}, L) := R^1 q'_* (j_*(p^{o,*}\mathcal{F} \otimes L))$$

for a local system  $\mathcal{F}$  on  $C = \overline{C} - D$ . Thus,  $MC(\mathcal{F}, L)$  is a local system on C too. Notice, in this formula, that  $j_*$  is not derived, indeed it is the middle extension that leads to the terminology for Katz's construction.

We say that a convolution kernel L is MC-invertible if there is another local system L' such that  $MC(MC(\mathcal{F}, L), L') \cong \mathcal{F}$ . We recall that this is the case for rank 1 local systems [41, 23].

**Remark 3.1.** If L is MC-invertible and if  $\mathcal{F}$  is irreducible then  $MC(\mathcal{F}, L)$  is irreducible.

If the convolution kernel is quadratic then the middle convolution operation preserves quadraticity:

**Proposition 3.2.** The dual local systems are related by the formula

$$MC(\mathcal{F}, L)^{\vee} \cong MC(\mathcal{F}^{\vee}, L^{\vee}).$$

In particular, if  $\mathcal{F}$  and L are quadratic then so is  $MC(\mathcal{F}, L)$ . The operation  $\mathcal{F} \mapsto MC(\mathcal{F}, L)$  will be called a quadratic middle convolution.

*Proof.* This was noted for rank 1 convolution kernels in [26, Remark 3.1.3]. The same holds in general, indeed middle convolution is obtained by composing together several operations on perverse sheaves over  $\overline{C} = \mathbb{P}^1$  or  $\overline{X}$  that commute with Verdier duality.

Katz's book [41] and several subsequent references were mainly about the case when the convolution kernel has rank 1. The case of higher rank convolution kernels was of course envisioned in [41] and was studied for example in [26, 48, 49, 25].

We start by looking at the rank 1 case, but in order to get an operation preserving quadraticity and acting on the eigenspaces of eigenvalues different from  $\pm 1$ , it will be useful to look at convolution kernels of rank 2 coming from classical hypergeometric local systems.

3.1. Rank 1 convolution kernels. The reductions using middle convolution in our case will pass through local systems with non-semisimple monodromy, even if we start with semisimple monodromy. Therefore, we need notation for the Katz middle convolution, involving convolution kernels Lof rank 1, for local systems with general monodromy conjugacy classes as was originally done by Katz [41], see also Dettweiler-Reiter [22, 23, 24].

Fix an integer r, the rank of the local system. A conjugacy class in GL(r)is determined by the following data: a partition  $r = \lambda_1 + \ldots + \lambda_k$  denoted  $\lambda$ , together with a collection of complex numbers  $a_1, \ldots, a_k$ , up to simultaneous permutations of the two. No condition of distinctness of the  $a_i$  is made. This determines a Jordan normal form in the following way. For each eigenvalue a(one of the  $a_i$ 's), let  $e^0(a), e^1(a), \ldots, e^{k(a)}(a)$  be the collection of parts of the partition  $\lambda$  whose associated eigenvalue is a, ordered by  $e^0(a) \geq e^1(a) \geq \dots$ This is a partition of  $r(a) := e^{0}(a) + e^{1}(a) + \dots$  which is the rank of the a-generalized eigenspace. Take the dual partition of that, and use that to determine the sizes of the Jordan blocks. Thus,  $e^{j}(a)$  is the number of Jordan blocks of size  $\geq j$  for the eigenvalue a. We note that this is similar to Katz's notation in [41, Section 6.0.2].

The dimension of the conjugacy class in GL(r) is  $r^2 - 2 - \sum \lambda_i^2$ . So, this notation behaves as if we were speaking of semisimple conjugacy classes with the  $\lambda_i$  being the multiplicity of the eigenvalues.

The operation of middle convolution with a rank 1 local system also behaves as if the monodromy were semisimple, with the only caveat being that we choose the largest part of the partition corresponding to the distinguished eigenvalue at each point.

Suppose  $\mathcal{F}$  is a local system. Let  $\lambda(i) = (\lambda_1(i), \dots, \lambda_r(i))$  be the partition (some of the  $\lambda_i(i)$  may be zero) and  $a_1(i), \ldots, a_r(i)$  the eigenvalues of the monodromy of  $\mathcal{F}$  at  $t_i$ . These numbers take on arbitrary values in case  $\lambda_i(i) = 0$  but by "eigenvalue" we only mean the values  $a_i(i)$  for which  $\lambda_i(i) > 0.$ 

A rank 1 convolution kernel  $\beta$  is characterized by specifying the horizontal, vertical and diagonal eigenvalues  $\beta^{H_i}$ ,  $\beta^{V_i}$  and  $\beta^T$  in  $\mathbb{C}^*$ , subject to the conditions that the product of the  $\beta^{H_i}$ , times  $\beta^T$ , equals 1; and similarly the product of the  $\beta^{V_i}$ , times  $\beta^T$ , equals 1. We call  $\alpha_i := (\beta^{H_i})^{-1}$  the distinguished eigenvalue at the point  $t_i$ . Say

that the index j is distinguished at  $t_i$  if  $a_j(i) = \alpha_i$  and if  $\lambda_j(i)$  is the largest

(nonzero) part with this eigenvalue, with j being the minimal value in case of equalities between the largest parts. There is a unique distinguished index at each point for which  $\alpha_i$  is among the eigenvalues.

Let  $\delta_i := \lambda_j(i)$  in case j is distinguished at i, and set  $\delta_i := 0$  if  $\alpha_i$  is not an eigenvalue. Define

$$\delta = \delta(\mathcal{F}, \beta) := r - \sum_{i} \delta_{i}.$$

Then the middle convolution operation sends the local system  $\mathcal{F}$  with given partition with notations as above, to a new local system  $\mathcal{F}'$  of rank  $r' = r + \delta$  with new partition data determined as follows. We define a new partition by setting  $\lambda'_j(i) := \lambda_j(i)$  if j is not distinguished at i; and  $\lambda'_j(i) = \lambda_j(i) + \delta$  if j is distinguished at i. If there is no distinguished index at i then one notes that  $\delta \geq 0$ ; if  $\delta > 0$ , we introduce a new part of size  $\delta$  indexed by j = r + 1.

The eigenvalue associated to the distinguished part at  $t_i$  is  $\beta^{V_i}$ .

For any non-distinguished part j at  $t_i$ , if  $a_j(i)$  was the associated eigenvalue for  $\mathcal{F}$  then the new associated eigenvalue for  $\mathcal{F}'$  is

$$a_j'(i) = \beta^{H_i} \beta^{V_i} \beta^T a_j(i).$$

In the semisimple case and under the conventions in place in the exposition of [58] designed to guarantee that the monodromy of  $\mathcal{F}'$  is also semisimple, it is easy to see that the middle convolution operation satisfies the above transformation rules. In the general case, that is also true by [41] as we shall state in the next proposition. This process may split off the top part for the distinguished eigenvalue to give a part that no longer has the same eigenvalues as the other parts that were associated to the distinguished eigenvalue. And, on the other hand, it can bring together two parts that previously were associated to different eigenvalues.

We note that if  $\mathcal{F}$  is irreducible then the size of the new part at eigenvalue  $\beta^{V_i}$ , namely  $\lambda_j(i)$  for  $\alpha_i = (\beta^{H_i})^{-1}$ , is greater than or equal to the  $\lambda_u(i)$  for an eigenvalue  $a_u(i)$  such that  $\beta^{H_i}\beta^{V_i}\beta^T a_u(i) = \beta^{V_i}$ . Indeed the latter condition means that  $\beta^{H_i}\beta^T a_u(i) = 1$ . Say  $\lambda_v(i') = \delta_{i'}$  for  $i' \neq i$  are the largest parts associated to the eigenvalues  $a_v(i') = (\beta^{H_i})^{-1}$ . Then the product of the  $a_v(i')$  is  $\beta^T\beta^{H_i}$ , so the product of the  $a_v(i')$  with  $a_u(i)$  is 1. Irreducibility thus implies that

$$\lambda_u(i) + \sum_{i' \neq i} \lambda_v(i') \le (k-2)r.$$

We have

$$\delta = (k-2)r - \lambda_j(i) - \sum_{i' \neq i} \lambda_v(i')$$

so  $\delta + \lambda_j(i) = (k-2)r - \sum_{i'\neq i} \lambda_v(i') \geq \lambda_u(i)$ . This proves the statement claimed at the start of this paragraph.

That means that when we add the new part as described above, it automatically goes to the first place in the ordered list of parts having the eigenvalue  $\beta^{V_i}$ .

**Proposition 3.3.** The above process describes the modification of conjugacy classes in all cases for which the rank 1 middle convolution is defined, that is to say whenever  $\beta^T \neq 1$ .

*Proof.* This is seen from the summary of the effect of Katz's operation [41, Sections 6.0.13, 6.1.7]. Dettweiler and Reiter state that the same formulation holds for their version of the middle convolution which is the one we are using here.

The following principle introduced in [41] is recorded here since it will be used often in the description of middle convolution operations.

**Principle 3.4.** Suppose  $a_i$  are distinguished eigenvalues for an irreducible rank  $r \geq 2$  local system  $\mathcal{F}$ , of multiplicities  $\lambda_i$  at the points  $t_1, \ldots, t_k$ . Suppose  $\delta := (k-2)r - \lambda_1 - \ldots - \lambda_k$  satisfies  $\delta < 0$ . Then for any rank 1 convolution kernel L with  $\beta^{H_i} = a_i^{-1}$ , we have  $\beta^T \neq 1$  so the middle convolution with L is defined.

Proof. If  $\beta^T = 1$  then there is a rank 1 local system M on  $\mathbb{P}^1 - \{t_1, \dots, t_k\}$  with eigenvalues  $a_i^{-1}$  at the points  $t_i$ . The Euler characteristic of the middle extension of  $M^{\vee} \otimes \mathcal{F}$  is at least  $-\delta > 0$ . Thus, either  $h^0(M^{\vee} \otimes \mathcal{F}) > 0$  in which case there is a nonzero map  $M \to \mathcal{F}$ , or else  $h_c^2(M^{\vee} \otimes \mathcal{F}) > 0$  in which case, by duality, there is a nonzero map  $\mathcal{F} \to M$ . Either one of these is a contradiction to irreducibility of  $\mathcal{F}$ .

3.2. Hypergeometric convolution kernels. A main observation leading to the reduction steps in the next section, is that the composition of two middle convolutions with rank 1 convolution kernels will in general be a middle convolution with a higher rank convolution kernel. Notice that the composition formulas in [41, 23, 24] concern a specific choice of horizontal and vertical eigenvalues in which case the composition corresponds again to a rank 1 convolution kernel. In the present section we envision a case, sufficient for our purposes, where the composition corresponds to a convolution kernel of rank 2 coming from a classical rank 2 hypergeometric local system on  $\mathbb{P}^1 - \{0, 1, \infty\}$ .

In this section, the assumption k = 3 is in effect.

Suppose U is a rank 2 hypergeometric local system over  $P := \mathbb{P}^1 - \{0, 1, \infty\}$  such that the monodromy transformation over 1 is the identity plus a matrix of rank 1. For the quadratic case this it is either unipotent or it has eigenvalues  $\pm 1$ .

Pull it back by a map  $f: X \to P$  to get the rank 2 local system  $L = f^*(U)$ . Use the map f, extending to the blown-up compactification  $\overline{X}$ , that sends the divisors  $H_1$  and  $V_2$  to 0,  $V_1$  and  $H_2$  to  $\infty$ , and T to 1. Note that L is trivial along  $H_3$  and  $V_3$ .

Along T, it has monodromy transformation that is the identity plus a rank 1 matrix.

The rank of  $\mathcal{F}'$  is obtained by calculating the middle cohomology of the fiber, that being the tensor product of U (moved to have singularities at  $t_1$ ,  $t_2$  and the diagonally moving point) with  $\mathcal{F}$ . The rank of cohomology on the open curve is 4r, the factor of 4 comes from two for the cohomology on  $\mathbb{P}^1$  minus four points, times two for the rank of U. The rank of the middle cohomology is then obtained by subtracting the dimensions of the local fixed parts.

The fixed part at the diagonal point has dimension r, indeed it comes by tensoring the nontrivial unipotent rank 2 piece by a constant vector space of dimension r and this has r-dimensional spaces of fixed or cofixed vectors.

The fixed part at  $t_3$  is just the fixed part in the monodromy of  $\mathcal{F}$ . Denote by a, a' and b, b' the eigenvalues of the monodromy of U at 0 and  $\infty$ , thus the eigenvalues of L along  $H_1$  are a, a' and the eigenvalues of L along  $H_2$ are b, b'. Then the fixed part at  $t_1$  comes from the  $(a)^{-1}$ -eigenspace and the  $(a')^{-1}$ -eigenspace of the monodromy of  $\mathcal{F}$  at  $t_1$ , and similarly using  $(b)^{-1}$ and  $(b')^{-1}$  at  $t_2$ .

Introduce the following notation:

$$\delta(L,\mathcal{F}) := r(\mathcal{F}) - r(MC(\mathcal{F},L)).$$

At each point  $t_i$  there is a set of distinguished eigenvalues, defined as the complex numbers  $\alpha$  such that  $\delta(L,\mathcal{F})$  is different from  $\delta(L,\mathcal{F}\otimes A)$  for a generic rank 1 local system, when  $\mathcal{F}$  has  $\alpha$  among its eigenvalues. The distinguished eigenvalues are thus invariants of the middle convolution operation MC(-,L). We see from the previous description that if  $L=f^*(U)$ with U having monodromy the identity plus a rank 1 matrix at 1, and having eigenvalues a, a' at 0 and b, b' at  $\infty$ , then the distinguished eigenvalues of MC(-, L) are  $(a)^{-1}$  and  $(a')^{-1}$  at  $t_1$ ,  $(b)^{-1}$  and  $(b')^{-1}$  at  $t_2$ , and 1 at  $t_3$ .

Lemma 3.5. In the above setting, the sets of distinguished eigenvalues of the operation MC(-,L) determine the hypergeometric local system U and hence the convolution kernel  $L = f^*(U)$ .

*Proof.* A classical rank 2 hypergeometric system is determined by its local monodromy transformations. It is therefore determined by its sets of local monodromy eigenvalues, since the local monodromy has to have a nontrivial unipotent part in case the two eigenvalues are equal. We have seen above that the local monodromy eigenvalues of U are the inverses of the distinguished eigenvalues of MC(-, L) at  $t_1$  and  $t_2$ .

**Proposition 3.6.** Consider the composition of two middle convolution operations MC(-, M) and MC(-, N) where M and N are rank 1 convolution kernels, having trivial eigenvalues at one of the points say  $t_3$ . In other words,  $eta_M^{H_3} = eta_M^{V_3} = eta_N^{H_3} = eta_N^{V_3} = 1.$ One possibility is that the composition is the identity up to tensoring with

a rank 1 system. Otherwise, there is either L of rank 1 or a classical rank

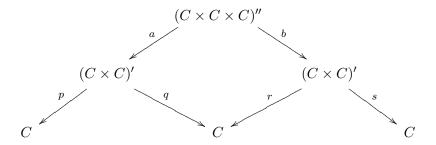
2 hypergeometric local system U leading to the pullback convolution kernel  $L = f^*U$  defined as above, such that

$$MC(MC(\mathcal{F}, M), N) \cong MC(\mathcal{F}, L) \otimes \Sigma$$

where  $\Sigma$  is a rank 1 local system on  $\mathbb{P}^1 - \{t_1, t_2\}$ .

The local monodromy transformations of U, and hence the isomorphism class of U, are determined by the effect of the composed middle convolution on the ranks and local monodromy data. If the local data vary in a way compatible with quadraticity, then L was quadratic and  $\Sigma$  has eigenvalues  $\pm 1$ .

*Proof.* Consider the diagram



where  $C = \mathbb{P}^1 - \{t_1, t_2, t_3\}$ ,  $(C \times C)'$  is the complement of the diagonal in  $C \times C$ , and  $(C \times C \times C)''$  is the complement of the big diagonal in  $C \times C \times C$ .

The two convolution kernels are rank 1 local systems M and N on  $(C \times C)'$ . The composition of the two middle convolution operators is obtained by taking  $\mathcal{F}$ , pulling back by p, tensoring with M, taking middle push-forward by q, pulling back by r, tensoring by N, and taking middle push-forward by s. Do the pullbacks, middle pushforwards, and tensor products in the category of perverse sheaves on  $\overline{C} = \mathbb{P}^1$ ,  $\overline{C} \times \overline{C}$  and  $\overline{C} \times \overline{C} \times \overline{C}$ .

By comparing with the middle extensions of the local systems on big open subsets, the above process is the same as pulling back by pa to  $C \times C \times C$ , tensoring with  $a^*(M) \otimes b^*(N)$ , and doing a middle pushforward by sb to C. On the other hand, we may factor pa and sb into a diagram of the form

$$(C \times C \times C)''$$

$$\downarrow pa \qquad \qquad w \qquad \qquad sb$$

$$C \stackrel{u}{\longleftarrow} (C \times C)' \stackrel{v}{\longrightarrow} C$$

where the middle vertical arrow is projection forgetting the middle factor. The composed operation described above is the same as first pulling back  $\mathcal{F}$  along u, tensoring with K, and pushing forward along v, where the convolution kernel is

$$K = R^1 w_{!*} (a^*(M) \otimes b^*(N)).$$

The fiber of w over a general point  $(x,z) \in C \times C$  is the curve  $C_{x,z} = \mathbb{P}^1 - \{t_1, t_2, t_3, x, z\}$ . The fiber of K over (x,z) is therefore the middle first cohomology of  $C_{x,z}$  with coefficients in  $a^*(M) \otimes b^*(N)$  (by which we mean this restricted to the fiber in question, a notation that will be left out in what follows). This latter is the tensor product of the restrictions  $M_x$  and  $N_z$ .

Let  $\alpha_1, \alpha_2$  and  $\alpha_3 = 1$  be the inverse horizontal eigenvalues of M, that is  $\beta_M^{H_i} = (\alpha_i)^{-1}$ . The vertical eigenvalues of M combine with the horizontal ones of N so we may as well assume that  $\beta_M^{V_i} = (\alpha_i)^{-1}$  too. We have  $\beta_M^T = \alpha_1 \alpha_2$ .

Similarly let  $\psi_1, \psi_2$  and  $\psi_3 = 1$  be the inverse horizontal eigenvalues of N so that  $\beta_N^{H_i} = (\psi_i)^{-1}$ . Again, the vertical eigenvalues of N may be absorbed into  $\Sigma$  so we can say  $\beta_N^{V_i} = (\psi_i)^{-1}$ , and  $\beta_N^T = \psi_1 \psi_2$ .

into  $\Sigma$  so we can say  $\beta_N^{V_i} = (\psi_i)^{-1}$ , and  $\beta_N^T = \psi_1 \psi_2$ . Now  $M_x$  has eigenvalues  $(\alpha_i)^{-1}$  over  $t_1, t_2$  and 1 over  $t_3$ , with eigenvalue  $\alpha_1 \alpha_2$  over x. Similarly,  $N_z$  has eigenvalues  $(\psi_1)^{-1}$  and  $(\psi_2)^{-1}$  over  $t_1, t_2$  but  $\psi_3 = 1$  again over  $t_3$ , and  $\psi_1 \psi_2$  over  $t_3$ .

The rank 1 local system  $a^*(M) \otimes b^*(N)$  restricted to the fiber  $C_{x,z}$  has trivial monodromy at the point  $t_3$ . Thus, the point  $t_3$  may now be ignored for the middle direct image of  $a^*(M) \otimes b^*(N)$  by w, so this middle direct image has rank 2 and we are (usually—see below) in the classical picture of the hypergeometric rank 2 local system obtained as a higher direct image of a rank 1 local system along a map with four singular points in the fibers. The monodromy at  $t_1$  is  $(\alpha_1)^{-1}(\psi_1)^{-1}$ , the monodromy at  $t_2$  is  $(\alpha_2)^{-1}(\psi_2)^{-1}$ , the monodromy at  $t_3$  is  $t_4$  is  $t_4$  and the monodromy at  $t_4$  is  $t_4$  is  $t_4$  and the monodromy at  $t_4$  is  $t_4$  is  $t_4$  and the monodromy at  $t_4$  is  $t_4$  is required.

Existence of MC(-,M) and MC(-,N) contains implicitly the assumptions that  $\alpha_1\alpha_2\neq 1$  and  $\psi_1\psi_2\neq 1$ . If both  $\alpha_1\psi_1$  and  $\alpha_2\psi_2$  were equal to 1 then it would be the case considered by [23] of middle convolution by inverse convolution kernels, where the composition is the identity. Otherwise, if one of these is equal to 1 then the higher direct image has rank 1 and we are in the case where U has rank 1. We may therefore assume that all four of the eigenvalues of  $a^*(M)\otimes b^*(N)$  are nontrivial, yielding the hypergeometric picture.

Pursuing the computation would involve looking at how the hypergeometric system is tensored with rank 1 systems on the input and output sides. This is left as an exercise.

Instead, by Lemma 3.5, the eigenvalues and hence the isomorphism class of the hypergeometric local system U are determined by the effect of  $\mathcal{F} \mapsto MC(MC(\mathcal{F}, M), N)$  on local monodromy data.

If the effect on local data is compatible with quadraticity, then the sets of distinguished eigenvalues of MC(-,L) are invariant under  $z\mapsto z^{-1}$ , and this implies that L is quadratic. The operation MC(-,L) therefore preserves quadraticity. The eigenvalues of  $\Sigma$  are seen from the effect on local monodromy data, so if the effect of the composition of MC(-,M) and

MC(-,N) is compatible with quadraticity then  $-\otimes \Sigma$  must be compatible with quadraticity, implying that the eigenvalues of  $\Sigma$  are  $\pm 1$ .

Corollary 3.7. A pullback rank 2 local system  $L = f^*(U)$  obtained from a classical rank 2 hypergeometric U, is an MC-invertible convolution kernel.

*Proof.* Any such L occurs in the situation of the proposition, for an appropriate choice of M, N. Each rank 1 convolution kernel is MC-invertible, so the composition of the two middle convolution operations is invertible.  $\square$ 

#### 4. The main reductions

From here on, the number of singular points is k = 3.

**Lemma 4.1.** Suppose given an irreducible MC-minimal quadratic local system  $\mathcal{F}$  of rank r > 2. Then for the multiplicities  $m_i^j$  of paired eigenvalues, i.e. those different from  $\pm 1$ , we have

$$m_1^j + m_2^j + m_3^j \le r.$$

*Proof.* Let  $m_i^0$  be the maximum of the multiplicities  $m_i^j$  at each point. Let  $a_i^0$  and  $(a_i^0)^{-1}$  be the paired eigenvalues of multiplicity  $m_i^0$  over the point  $t_i$ ,  $a_i^0 \neq \pm 1$ . Set  $\delta := r - m_1^0 - m_2^0 - m_3^0$ . We aim to show that  $\delta \geq 0$  so let us assume  $\delta < 0$  and derive a contradiction. By Principle 3.4,  $a_1^0 a_2^0 a_3^0 \neq 1$ .

Apply Katz's middle convolution with L a rank 1 local system having eigenvalues  $\beta^{H_i} = \beta^{V_i} := (a_i^0)^{-1}$  and  $\beta^T = a_1^0 a_2^0 a_3^0 \neq 1$ . The resulting local system  $\mathcal{F}'$  has eigenvalues of multiplicities  $m_i^0 + \delta$ , they are  $\beta^{V_i} = (a_i^0)^{-1}$ , coming from the eigenvalues  $a_i^0$  in the original local system. However, the original local system also has eigenvalues  $(a_i^0)^{-1}$  of multiplicity  $m_i$  and these lead to eigenvalues

$$\phi_i := \beta^{H_i} \beta^T \beta^{V_i} (a_i^0)^{-1} = (a_1^0 a_2^0 a_3^0) (a_i^0)^{-3}$$

that remain of multiplicity  $m_i^0$ . Middle convolution with L is invertible hence preserves irreducibility.

However, we note that  $\phi_1\phi_2\phi_3 = 1$ . The new local system  $\mathcal{F}'$  has rank  $r + \delta$ . There is a rank one local system U with eigenvalues  $\phi_1, \phi_2, \phi_3$ . Now  $\mathcal{F}' \otimes U^{\vee}$  has trivial eigenvalues of multiplicities  $m_i$  at the points  $t_i$ , but

$$m_1 + m_2 + m_3 > r > r + \delta = r(\mathcal{F}').$$

As in the proof of Principle 3.4, this contradicts the irreducibility of  $\mathcal{F}'$  unless the rank  $r+\delta=1$ . In turn, that could only happen if  $m_i+\delta=0$  and  $m_i=1$  and there are no other eigenvalues, in other words if  $\mathcal{F}$  was a hypergeometric rank 2 local system. This completes the proof of a contradiction unless r=2.

An alternate proof of the lemma would be to say directly that if  $\delta < 0$  then there has to be a nontrivial morphism from the hypergeometric rank 2 local system with eigenvalues  $a_i^0$  and  $(a_i^0)^{-1}$  at  $t_i$ , to  $\mathcal{F}$ .

**Lemma 4.2.** Suppose given an irreducible MC-minimal quadratic local system  $\mathcal{F}$  of rank > 1. Then for the highest half-unital eigenvalue multiplicities  $e_i^0$  we have

$$e_1^0 + e_2^0 + e_3^0 \le r$$
.

Proof. Let  $\sigma_i$  be the half-unital eigenvalues with parts  $e_i^0$ , so  $\sigma_i = \pm 1$ . If there are an even number of  $\sigma_i = -1$ , we may assume that all  $\sigma_i = 1$  by tensoring with a rank 1 local system having monodromy -1 at two points, if necessary. Then, if  $e_1^0 + e_2^0 + e_3^0 > r$  it follows that  $h^0(\mathcal{F}) > 0$ , contradicting irreducibility. This shows the inequality when the number of negative  $\sigma_i$  is even

Assume the number of negative  $\sigma_i$  is odd and the inequality in question doesn't hold. Tensoring with a rank 1 system we may assume that  $\sigma_i = -1$  for i = 1, 2, 3. Let L be a rank 1 convolution kernel with  $\beta^{H_i} = \beta^{V_i} = \beta^T = -1$ , and apply middle convolution with L. The rank changes by

$$\delta = r - e_1^0 + e_2^0 + e_3^0 < 0.$$

The local system L is quadratic so  $\mathcal{F}' = MC(\mathcal{F}, L)$  is again quadratic. The new rank  $r' = r + \delta$  is strictly smaller. The rank 1 middle convolutions are invertible [41, 23]. This contradicts MC-minimality of  $\mathcal{F}$ , completing the proof.

For information, we look at how the partitions change under this operation, as an illustration of the general setup. The parts of size  $e_i^0$  become parts associated again to eigenvalues  $\beta^{V_i}=-1$ , with multiplicities  $e_i^0+\delta$ . Note that  $\beta^{H_i}\beta^{V_i}\beta^T=-1$  so the remaining parts  $e_0^j$  (j>0) that were associated to the eigenvalue  $\sigma_i=-1$  become associated to eigenvalue 1. The parts  $f_i^j$  associated to eigenvalue 1 become associated to -1. We note that  $e_1^0+e_2^0+f_3^0\leq r$  by the argument of the first paragraph, and permutations of this. It follows that  $f_i^0\leq e_i^0+\delta$  so the new parts go on top of the  $f_i^0$ .

Notice that even if the monodromy transformations were semisimple before the middle convolution operation, they will in general have nontrivial unipotent parts after the operation.  $\Box$ 

The dimension remains the same under the quadratic middle convolution operation of the previous lemma, since it is invertible, but it would be interesting to verify that combinatorially. In particular, it is strongly suggested that the type  $\varepsilon$  changes under this operation, although we didn't do a separate verification of that.

**Lemma 4.3.** Suppose given an irreducible MC-minimal quadratic local system  $\mathcal{F}$ . Then for the paired eigenvalue multiplicities  $m_i^j$  and the highest half-unital eigenvalue multiplicities  $e_i^0$  we have

$$m_1^j + m_2^j + e_3^0 \le r,$$

plus permutations of this inequality.

*Proof.* We may assume that the half-unital eigenvalue with multiplicity  $e_3^0$  over  $t_3$  is 1. Indeed if it were -1 we could modify it by tensoring with a rank 1 local system with eigenvalues -1 at two points say  $t_1, t_3$ .

Let  $m_i^0$  be the maximal paired multiplicities, and let  $\delta := r - m_1^0 - m_2^0 - e_3^0$ , we would like to show  $\delta \geq 0$  so we may assume  $\delta < 0$ . Let  $a_i^0$  and  $(a_i^0)^{-1}$  be the paired eigenvalues of multiplicity  $m_i^0$  over the points  $t_1, t_2$ .

We are going to make a reduction by middle convolution. There are two approaches related by Proposition 3.6.

Look first at what happens upon doing two Katz middle convolutions by rank 1 convolution kernels that are however not individually quadratic. At the first step, use a local system M with eigenvalues  $\beta_M^{H_i} = \beta_M^{V_i} = (a_i^0)^{-1}$  for i=1,2 and trivial eigenvalues for i=3, with  $\beta_M^T = a_1^0 a_2^0$ . As usual, by Principle 3.4 we may assume that  $\beta_M^T \neq 1$  because  $\mathcal{F}$  is assumed irreducible.

The resulting local system  $\mathcal{F}' = MC(\mathcal{F}, M)$  has rank  $r' = r + \delta$ . The eigenvalues  $a_i^0$  become  $(a_i^0)^{-1}$  with multiplicities  $m_i^0 + \delta$ . The half-unital eigenvalue 1 at i = 3 stays as  $\phi_3 = 1$  (since that was the corresponding  $\beta_M^{V_3}$ ) with multiplicity  $e_3^0 + \delta$ . A half-unital eigenvalue -1 over  $t_3$  becomes  $\varphi_3 = -a_1^0 a_2^0$  with multiplicity staying at  $f_3^0$ . The eigenvalues  $(a_i^0)^{-1}$  become

$$\phi_i = (a_1^0 a_2^0)(a_i^0)^{-2}(a_i^0)^{-1} = (a_1^0 a_2^0)(a_i^0)^{-3}$$

with the same multiplicities  $m_i^0$ .

We have

$$\phi_1 = a_2^0/(a_1^0)^2, \quad \phi_2 = a_1^0/(a_2^0)^2.$$

Notice that the situation here is different from that of the previous proof: we have  $\phi_1\phi_2\phi_3 = (a_1^0a_2^0)^{-1} \neq 1$ , because the eigenvalue  $\phi_3$  was modified according to the formula for the eigenvalue whose multiplicity changes, rather than the formula for one whose multiplicity stays the same.

We may therefore do a second middle convolution  $\mathcal{F}'' = MC(\mathcal{F}', N)$  with distinguished eigenvalues  $\phi_i$  of multiplicities  $m_1^0, m_2^0, e_3 + \delta$ . The new invariant is

$$\delta' = r' - m_1^0 - m_2^0 - (e_3 + \delta) = r + \delta - m_1^0 - m_2^0 - (e_3 + \delta) = \delta.$$

The diagonal eigenvalue of the new operation is  $\beta_N^T = \phi_1\phi_2\phi_3 = (a_1^0a_2^0)^{-1}$ , whereas the horizontal eigenvalues are  $\beta_N^{H_i} = \phi_i^{-1}$  for i=1,2 and trivial for i=3. Let us use as vertical eigenvalues here  $\beta_N^{V_i} = \xi_i := \phi_i(a_i^0)^2 = (a_1^0a_2^0)(a_i^0)^{-1}$ , (in other words  $\xi_1 = a_2^0$  and vice-versa) noting that  $\xi_1\xi_2$  is still the inverse of the diagonal eigenvalue. Take the vertical eigenvalue to be trivial at i=3.

Over  $t_3$  the eigenvalue 1 of multiplicity  $e_3+\delta$  becomes again 1 of multiplicity  $e_3+2\delta$ . The eigenvalue  $\varphi_3$  of multiplicities  $f_3^j$  becomes  $\varphi_3 \cdot (\phi_1 \phi_2 \phi_3) = -1$  again of multiplicity  $f_3^j$ . Similarly, the other parts of multiplicity  $e_3^j$  with  $j \geq 1$  return to 1.

Over  $t_1, t_2$  the eigenvalues change as follows. For the ones other than those coming from  $a_i^0$  and  $(a_i^0)^{-1}$ , we note that the products of the horizontal,

vertical and diagonal eigenvalues multiplied together for the two operations, are

$$(a_i^0)^{-2} \cdot (a_1^0 a_2^0) \cdot \phi_i^{-1} \cdot \xi_i \cdot (\phi_1 \phi_2) = 1$$

so the eigenvalues that were originally different from  $a_i^0$  and  $(a_i^0)^{-1}$  stay the same, with the same multiplicities. Furthermore, any parts that have eigenvalues  $a_i^0$  or  $(a_i^0)^{-1}$  but that correspond to lower pieces in the partitions at those eigenvalues, stay the same with the same multiplicities.

The eigenvalue  $a_i^0$  becomes first  $(a_i^0)^{-1}$  and then

$$(a_i^0)^{-1} \cdot \phi_i^{-1} \cdot \xi_i \cdot (\phi_1 \phi_2) = (a_i^0) \cdot (a_1^0 a_2^0)^{-1} = \xi_i^{-1}$$

with multiplicity  $m_i^0 + \delta$ . The eigenvalue  $(a_i^0)^{-1}$  becomes first  $\phi_i$  of multiplicity  $m_i^0$  which is then transformed to the vertical eigenvalue  $\xi_i$  with multiplicity  $m_i^0 + \delta$ . Note in particular that these are again paired eigenvalues, we have in fact just interchanged the eigenvalues of the points  $t_1$  and  $t_2$  and changed the multiplicities.

The second middle convolution operation yields a local system  $\mathcal{F}''$  of rank  $r'' = r + 2\delta$ . The new multiplicities are again  $m_i^0 + \delta$  so at each point we now have two eigenspaces of multiplicities  $m_i^0 + \delta$ . The paired eigenvalues are again inverses and the ones of multiplicities  $e_3^0 + 2\delta$  and  $e_3^j$   $(j \geq 1)$  and  $f_3^j$  remain half-unital. This composition of two middle convolutions therefore respects the quadraticity condition on the local monodromy data.

As in Proposition 3.6, this composition of two middle convolutions may be viewed as middle convolution with a single hypergeometric local system. Let U be the rank 2 hypergeometric local system over  $P := \mathbb{P}^1 - \{0, 1, \infty\}$  having monodromy transformations semisimple with eigenvalues  $a_1^0$  and  $(a_1^0)^{-1}$  at 0, semisimple with eigenvalues  $a_2^0$  and  $(a_2^0)^{-1}$  at  $\infty$ , and nontrivial unipotent over 1. Pull this back by a map  $X \to P$  to get the rank 2 local system L which is trivial over  $V_3$  and  $H_3$ . Along T, it has nontrivial unipotent monodromy. Over  $V_1$  and  $H_1$  it has semisimple monodromy with eigenvalues  $a_1^0$  and  $(a_1^0)^{-1}$ , and over  $V_2$  and  $H_2$  it has semisimple monodromy with eigenvalues  $a_2^0$  and  $(a_2^0)^{-1}$ .

Notice that U and hence L are self-dual, since a hypergeometric local system is determined by its local monodromy eigenvalues. Thus, middle convolution with L sends a quadratic local system  $\mathcal{F}$  to another quadratic local system  $\mathcal{F}'' = MC(\mathcal{F}, L)$ . By the method used in the proof of Proposition 3.6, this is the same as the previous  $\mathcal{F}''$ . Indeed, the distinguished eigenvalues for MC(-, L) are the same as those of the combined operation MC(MC(-, M), N). This may be seen as in the next paragraphs.

One may calculate directly the rank of  $\mathcal{F}''$  from middle convolution with L. This redoes a part of the proof of Proposition 3.6 but it is useful to see how it works in the context of this application, notably in order to show that the specification of eigenvalues of L given above was the right one.

The new rank is obtained by calculating the middle cohomology of the fiber, that being the tensor product of U (moved to have singularities at

 $t_1, t_2$  and the diagonally moving point) with  $\mathcal{F}$ . In turn the rank of the middle cohomology is obtained by subtracting the dimensions of the local fixed parts from the rank 4r of cohomology on the open curve. The fixed part at the diagonal point has dimension r, indeed it comes by tensoring the nontrivial unipotent rank 2 piece by a constant vector space of dimension r and this has r-dimensional spaces of fixed or cofixed vectors. The fixed part at  $t_1$  consists of the piece of dimension  $m_1^0$  obtained by tensoring the  $a_1^0$ -eigenspace of  $\mathcal{F}$  with the  $(a_1^0)^{-1}$ -eigenspace of L, plus another piece of dimension  $m_1^0$  obtained from the dual eigenspaces. Thus the fixed part at  $t_1$  has dimension  $2m_1^0$ . Similarly the fixed part at  $t_2$  has dimension  $2m_2^0$ . Finally, at the point  $t_3$  the local system L is trivial of rank 2, tensoring with  $\mathcal{F}$  that has an  $e_3^0$ -dimensional eigenspace of eigenvalue 1, so this has fixed part of dimension  $2e_3^0$ . Altogether we get

$$r' = r(MC(\mathcal{F}, L)) = 4r - r - 2m_1^0 - 2m_2^0 - 2e_3 = r + 2\delta$$

with  $\delta := r - m_1^0 - m_2^0 - e_3$  as above. Under our hypothesis  $\delta > 0$  we get to a smaller rank quadratic local system  $\mathcal{F}''$ . By Corollary 3.7, L is MC-invertible. Thus, the decrease in rank contradicts the hypothesis that  $\mathcal{F}$  is MC-minimal, completing the proof.

Lemma 4.4. Suppose given an MC-minimal irreducible quadratic local system  $\mathcal{F}$  of rank r > 2. Then for the paired eigenvalue multiplicities  $m_i^j$  (with  $m_i^0$  being the biggest) and the half-unital parts  $e_i^0 \geq f_i^0$  we have

$$m_1^0 + e_2^0 + (e_3^0 + f_3^0)/2 \le r.$$

Similarly for permutations of this.

*Proof.* As before, assume that the inequality doesn't hold, and assume by tensoring if necessary that the parts  $e_2^j$  and  $e_3^j$  correspond to eigenvalue  $\sigma_2 = \sigma_3 = 1.$ 

We consider first a convolution kernel M with distinguished eigenvalues  $a_1^0$ , 1 and 1. Precisely,  $\beta_M^{H_1} = \beta_M^{V_1} = (a_1^0)^{-1}$ ,  $\beta_M^{H_2} = \beta_M^{V_2} = \beta_M^{H_3} = \beta_M^{V_3} = 1$ , and  $\beta_M^T = a_1^0 \neq 1$ . For this middle convolution operation,

$$\delta = r - m_1^0 - e_2^0 - e_3^0.$$

The product  $\beta_M^{H_i}\beta_M^{V_i}\beta_M^T$  is equal to  $(a_1^0)^{-1}$  for i=1, and equal to  $a_1^0$  for

This sends  $\mathcal{F}$  to a new local system (no longer quadratic)  $\mathcal{F}'$  of rank  $r' = r + \delta$  with the following parts.

Over  $t_1$  we have a part of size  $m_1^0 + \delta$  associated to  $\beta_M^{V_1} = (a_1^0)^{-1}$ , a part of size  $m_1^0$  associated to  $\beta_M^{H_1} \beta_M^{V_1} \beta_M^{T} (a_1^0)^{-1} = (a_1^0)^{-2}$ , plus the remaining parts of the same sizes as before associated to  $(a_1^0)^{-1}$  times the previous eigenvalues. Notice that if 1 was among the previous eigenvalues then the associated parts are added to the new one at the eigenvalue  $(a_1^0)^{-1}$ , with the remarks from prior to Proposition 3.3 applying to say that the new part goes on top.

Over  $t_2$  we get a part of size  $e_2^0 + \delta$  associated to the eigenvalue  $\beta_M^{V_2} = 1$ , plus the other parts associated to their eigenvalues multiplied by  $a_1^0$ .

Similarly, over  $t_3$  we get a part of size  $e_2^0 + \delta$  associated to the eigenvalue  $\beta_M^{V_2} = 1$ , plus the other parts associated to their eigenvalues multiplied by  $a_1^0$ . In particular, this includes a part of size  $f_3^0$  associated to the eigenvalue  $-a_1^0$ .

Next do a middle convolution with distinguished eigenvalue  $(a_1^0)^{-2}$  over  $t_1$  associated to a part of size  $m_1^0$ ; distinguished eigenvalue 1 associated to the part of size  $e_2^0 + \delta$  over  $t_1$ ; and distinguished eigenvalue  $-a_1^0$  associated to the parts of size  $f_3^0$ . However, we tensor appropriately the eigenvalues after the convolution, to the effect that we use a convolution kernel N whose eigenvalues are

$$\beta_N^{H_1}=(a_1^0)^2, \ \ \beta_N^{V_1}=1, \ \ \beta_N^{H_2}=1, \ \ \beta_N^{V_2}=1, \ \ \beta_N^{H_3}=-(a_1^0)^{-1}, \ \ \beta_N^{V_3}=-a_1^0,$$
 and  $\beta_N^T=-(a_1^0)^{-1}\neq 1.$  We get

$$\beta_N^{H_1}\beta_N^{V_1}\beta_N^T = -a_1^0, \quad \beta_N^{H_2}\beta_N^{V_2}\beta_N^T = -(a_1^0)^{-1} \quad \beta_N^{H_3}\beta_N^{V_3}\beta_N^T = -(a_1^0)^{-1}.$$

The change of rank is given by

$$\delta' = r' - m_1^0 - (e_2^0 + \delta) - f_3^0 = r - m_1^0 - e_2^0 - f_3^0 = \delta + e_3^0 - f_3^0.$$

The resulting local system  $\mathcal{F}''$  has rank  $r'' = r' + \delta'$ . If we assume that the inequality to be proven in the lemma is not true, then r'' < r. Indeed,

$$\delta'' := \delta + \delta' = 2r - 2m_1^0 - 2e_2^0 - e_3^0 - f_3^0$$

which is twice the difference of two sides of our desired inequality.

Look at the local monodromy transformations for  $\mathcal{F}''$ . Over  $t_1$  we have a new part of size  $m_1^0 + \delta'$  with eigenvalue  $\beta_N^{V_i} = 1$ , and a part coming from the new part in  $\mathcal{F}'$ , of size  $m_1^0 + \delta$  with eigenvalue  $\beta_N^{H_1} \beta_N^{V_1} \beta_N^T (a_1^0)^{-1} = -1$ . There are also the parts whose size is unchanged, corresponding to the previous eigenvalues multiplied by -1. We note that these can combine with the new parts to create nontrivial unipotent conjugacy classes even if we started with semisimple ones. The previous paired eigenvalues remain paired.

Notice also here that the previous parts, corresponding to paired eigenvalues, have turned into parts corresponding to half-unital eigenvalues of sizes that are in general different since  $\delta$  and  $\delta'$  are in general different.

Over  $t_2$  the new part in  $\mathcal{F}''$  has size  $e_2^0 + \delta''$  and is associated to eigenvalue  $\beta_N^{V_2} = 1$ . The other parts are the ones coming from the original partition, with associated eigenvalues multiplied by -1. In particular, the parts of sizes  $e_2^j$  for  $j \geq 1$ , become associated to -1. The ones of sizes  $f_2^j$  become associated to eigenvalue 1 and are combined with the new part, with the new part going on top. The previous paired eigenvalues remain paired.

Over  $t_3$ , the new part has size  $f_3^0 + \delta'$ , associated to eigenvalue  $-a_1^0$ . The previous new part of size  $e_3^0 + \delta$  that had eigenvalue 1, is now associated to eigenvalue  $-(a_1^0)^{-1}$ . We note that  $f_3^0 + \delta' = e_3^0 + \delta$  and these two eigenvalues form a conjugate pair. Thus, the half-unital eigenvalues have

been transformed into paired ones. The other eigenvalues get multiplied by  $a_1^0 \cdot (-(a_1^0)^{-1}) = -1$ .

We see that the local monodromy transformations are compatible with a quadratic structure. Proceeding as in the proof of Lemma 4.3 using Proposition 3.6, we can interpret the composition of two middle convolutions as a middle convolution with a quadratic and MC-invertible rank 2 convolution kernel L coming from a hypergeometric system. The hypothesis  $\delta'' < 0$  implies that the new rank is strictly smaller than r, and this contradicts the hypothesis that  $\mathcal{F}$  was MC-minimal. This shows that  $\delta'' \geq 0$ , completing the proof.

**Lemma 4.5.** Suppose given an MC-minimal irreducible quadratic local system  $\mathcal{F}$  of rank r > 2. Then for the paired eigenvalue multiplicities  $m_i^j$  (with  $m_i^0$  being the biggest) and the half-unital parts  $e_i^0 \geq f_i^0$  we have

$$m_1^0 + e_2^0 + (e_3^0 + e_3^1)/2 \le r$$
.

Similarly for permutations of this.

*Proof.* Proceed as in the previous proof. We'll just indicate the changes here. The first step is the same. At the second step, we choose as distinguished parts the ones of sizes  $m_1^0$  over  $t_1$ , of size  $e_2^0 + \delta$  over  $t_2$ , and of size  $e_3^1$  over  $t_3$ . Notice that the eigenvalue associated to the part of size  $e_3^1$  is  $a_1^0$ .

Therefore, we use this time a convolution kernel N with

$$\beta_N^{H_1}=(a_1^0)^2, \ \beta_N^{V_1}=1, \ \beta_N^{H_2}=1, \ \beta_N^{V_2}=1, \ \beta_N^{H_3}=(a_1^0)^{-1}, \ \beta_N^{V_3}=a_1^0,$$
 and  $\beta_N^T=(a_1^0)^{-1}.$  We get

$$\beta_N^{H_1}\beta_N^{V_1}\beta_N^T=a_1^0,\quad \beta_N^{H_2}\beta_N^{V_2}\beta_N^T=(a_1^0)^{-1}\quad \beta_N^{H_3}\beta_N^{V_3}\beta_N^T=(a_1^0)^{-1}.$$

The change of rank is given by

$$\delta' = r' - m_1^0 - (e_2^0 + \delta) - e_3^1 = r - m_1^0 - e_2^0 - f_3^0 = \delta + e_3^0 - e_3^1.$$

The resulting local system  $\mathcal{F}''$  has rank  $r'' = r' + \delta'$ . If we assume that the inequality to be proven in the lemma is not true, then r'' < r. As before,

$$\delta'' := \delta + \delta' = 2r - 2m_1^0 - 2e_2^0 - e_3^0 - e_3^1$$

which is twice the difference of two sides of our desired inequality.

The discussion of local monodromy transformations is similar to the previous one, with the - signs removed. Again, it preserves the quadratic property of the local monodromy transformations, replacing the distinguished paired eigenvalues over  $t_1$  by half-unital parts, and replacing the dinstinguished half-unital eigenvalue parts over  $t_3$  by a two paired ones.

Again using Proposition 3.6, we can interpret the composition of two middle convolutions as a middle convolution with a quadratic and MC-invertible rank 2 convolution kernel coming from a hypergeometric system. If  $\delta'' < 0$  then the rank decreases, contradicting MC-minimality of  $\mathcal{F}$ , so this completes the proof.

4.1. Numerical MC-minimality. Recall that k = 3. A collection of partition data consists of the specification of  $m_i^j$ ,  $e_i^j$  and  $f_i^j$  for each i = 1, 2, 3 subject to the partition conditions that

$$2\sum_{j}m_{i}^{j}+\sum_{j}e_{i}^{j}+\sum_{j}f_{i}^{j}=r$$

(the three sums are not over the same sets of indices j), and subject to the order conditions  $e_i^0 \ge e_i^1 \ge \dots$  and  $f_i^0 \ge f_i^1 \ge \dots$ 

From these we let  $(\lambda_i^j)_j$  be the full collection of parts including the  $e_i^j$ , the  $f_i^j$  and each  $m_i^j$  counted twice, so  $r = \sum_j \lambda_i^j$ . Define a decoration to the partition consisting of  $\epsilon_i^j \in \{-1,0,1\}$  defined by  $\epsilon_i^j = 0$  for parts corresponding to  $m_i^j$ , and  $\epsilon_i^j = (-1)^u$  if  $\lambda_i^j$  corresponds to  $e_i^u$  or  $f_i^u$ .

Recall that  $\varepsilon = \pm 1$  is fixed depending on the group,  $\varepsilon = 1$  for the orthogonal group and  $\varepsilon = -1$  for the symplectic group. The *dimension* associated to the partition data is

$$\Delta = r^2 - \sum_{i,j} (\lambda_i^j)^2 - \varepsilon \left( r - \sum_{i,j} \epsilon_i^j \lambda_i^j \right).$$

This is in fact twice the complex dimension of the quadratic relative character variety, notation being used to simplify the factors of 1/2.

The term  $\beta := r^2 - \sum_{i,j} (\lambda_i^j)^2$  will be called the *box dimension* because it will correspond to the area left over after the collection of squares of size  $\lambda_i^j \times \lambda_i^j$  are fit into a square of size  $r \times r$ . The term

$$\ell := r - \sum_{i,j} \epsilon_i^j \lambda_i^j$$

will be called the *linear correction*, so the dimension is  $\Delta = \beta - \varepsilon \ell$ .

**Definition 4.6.** We say that (r, m; e; f;) is numerically MC-minimal if the inequalities of Lemmas 4.1, 4.2, 4.3, 4.4, 4.5 are satisfied.

**Proposition 4.7.** Any component of the relative character variety of quadratic local systems with trivial determinant is isomorphic, via middle convolution operations, to a component of the relative character variety for a numerically MC-minimal collection of decorated partition data.

*Proof.* Middle convolution with a quadratic MC-invertible convolution kernel induces an isomorphism between the corresponding quadratic relative character varieties. This is seen by using the expression of [23] for the middle convolution as an algebraic map on tuples of matrices. Then the combination of Lemmas 4.1, 4.2, 4.3, 4.4, 4.5, applied iteratively, gives the reduction to a numerically MC-minimal case.

**Remark 4.8.** The dimension of a component of the relative character variety is even, i.e.  $\Delta$  is a multiple of 4.

One can deform to unitary eigenvalues, then the relative character variety has a hyperkähler structure [38, 29, 30, 39, 52]. One can also give a combinatorial proof of this using the formula of Proposition 2.5 and the additional properties of unipotent conjugacy classes mentioned in Remark 2.2. That is left as an exercise.

We can now state the main theorem.

**Theorem 4.9.** Given  $d \geq 2$ , there is a constant R(d) such that for  $r \geq R(d)$  there are no numerically MC-minimal decorated partition data for rank r quadratic local systems with  $\Delta = 2d$ .

Using Proposition 4.7, this implies Theorem 1.1.

#### 5. Computations: General Considerations

Start now the proof of Theorem 4.9. The dimension d will be considered as fixed, and R is chosen successively large enough so that the various bounds hold for any numerically MC-minimal monodromy data of rank  $r \geq R$  and having  $\Delta = 2d$ .

Recall that  $e_i := e_i^0$  denotes the biggest of the half-unital (cf §2.2) eigenvalue parts at the point  $t_i$ , then  $f_i := f_i^0$  is the largest for the other half-unital eigenvalue. And  $m_i^j$  denote the multiplicities of the paired eigenvalues  $a_i^j \neq \pm 1$  and their duals  $(a_i^j)^{-1}$ . Let  $m_i^0$  be the biggest of these. Thus for each i,

$$r = e_i^0 + \ldots + f_i^0 + \ldots + 2m_i^0 + \ldots$$

#### 5.1. Fitting into a square.

**Proposition 5.1.** Any numerically MC-minimal collection of partitions  $(e_i^j, f_i^j, m_i^j)$  (Definition 4.6) has the property that the union of squares of sizes  $e_i^j \times e_i^j$ ,  $f_i^j \times f_i^j$  and  $m_i^j \times m_i^j$  fits into a square of size  $r \times r$  in one of two ways. In the non-overlapping case each collection of squares fits into a rectangle of size  $q_i \times r$  with  $q_1 + q_2 + q_3 = r$ . In the overlapping case (see Figure 1) there are two of the half-unital blocks that overlap; up to permutations of the points we may assume that these are  $e_2^0$  and  $e_3^0$ , then the squares for i = 1 fit into the rectangle of size  $q_1 \times r$  while the squares for i = 2 are either the  $e_2^0$  part, or fit into a rectangle of size  $q_2 \times r_2$  with  $r_2 = r - e_2^0$  and similarly the squares for i = 2 are either the  $e_3^0$  part, or fit into a rectangle of size  $q_3 \times r_3$  with  $r_3 = r - e_3^0$ . The  $e_2^0$  and  $e_3^0$  squares have an horizontal overlap of  $\mu \geq 1$  so that  $r = q_1 + e_2^0 + e_3^0 - \mu$ . We let  $\nu := r - e_2^0 - e_3^0$  (it is  $\geq 0$ ) and there is a leftover rectangle of size  $\mu \times \nu$  in the middle.

*Proof.* We have the inequalities

$$e_1^0 + e_2^0 + e_3^0 \le r$$
 
$$m_1^0 + m_2^0 + m_3^0 \le r$$
 
$$m_1^0 + m_2^0 + e_3^0 \le r$$
 and permutations.

In the case where  $m_i^0 \ge e_i^0$  for zero, two, or three of the columns, these lead to the non-overlapping case.

Assume therefore (up to a permutation of the *i*'s if necessary) that  $m_1^0 > e_1^0, e_2^0 > m_2^0$  and  $e_3^0 > m_3^0$ . Let

$$\mu := r - m_1^0 - e_2^0 - e_3^0.$$

If  $\mu \leq 0$  then we are back in the non-overlapping case, so assume  $\mu \geq 1$ . Set  $q_1 := m_1^0$ ,  $q_2 := e_2^0 - \mu$  and  $q_3 := e_3^0 - \mu$ . Set  $r_2 := r - e_2^0$  and  $r_3 := r - e_3^0$ . Divide the  $r \times r$  square into a rectangle of size  $q_1 \times r$ , the  $e_2^0 \times e_2^0$  square at the bottom with a rectangle of size  $q_2 \times r_2$  over it, and the  $e_3^0 \times e_3^0$  square at the top next to the previous rectangle, with the  $q_3 \times r_3$  rectangle below it. In the middle of all that is a rectangle of size  $\mu \times \nu$  with  $\nu = r - e_2^0 - e_3^0$ , note that  $\nu \geq 0$ .

The MC inequalities for two m's and an e imply that the squares of size  $m_i^j \times m_i^j$  fit into their corresponding rectangles of width  $q_i$ . Recall the additional MC inequalities

$$m_1^0 + e_2^0 + (e_3^0 + f_3^j)/2 \le r,$$

and similarly for  $f_2^j$ . These imply that the squares of sizes  $f_2^j \times f_2^j$  and  $f_3^j \times f_3^j$  also fit into the rectangles of size  $q_2 \times r_2$  and  $q_2 \times r_2$  respectively. Similarly the inequalities

$$m_1^0 + e_2^0 + (e_3^0 + e_3^j)/2 \le r \text{ (for } j \ge 1)$$

imply that for  $j \geq 1$ , the boxes of sizes  $e_2^j \times e_2^j$  and  $e_3^j \times e_3^j$  again fit into the rectangles of size  $q_2 \times r_2$  and  $q_2 \times r_2$  respectively. This is the overlapping configuration.

## 5.2. **Dimensions.** The box dimension is defined to be

$$\beta := r^2 - \sum_{i,j} (e_i^j)^2 - \sum_{i,j} (f_i^j)^2 - 2 \sum_{i,j} (m_i^j)^2.$$

In other words, it is the difference of the area of the big square minus all the smaller squares, noting that the multiplicities  $m_i^j$  each occur twice. This is the shaded area in Figure 1.

Define the linear correction to be

$$\ell := r - \sum_{i,j} (-1)^j (e_i^j + f_i^j).$$

Recall that  $\varepsilon = \pm 1$  determines whether we are looking at the orthogonal case ( $\varepsilon = 1$ ) or the symplectic case ( $\varepsilon = -1$ ). By Proposition 2.5, twice the dimension of the relative character variety at  $\mathcal{F}$  is

$$\Delta := 2 \dim \left( \mathscr{X}^{\text{irred}}(r, \varepsilon, C_1, \dots, C_k) \right) = \beta - \varepsilon \ell.$$

**Lemma 5.2.** The linear correction satisfies  $-r \le \ell \le r$ .

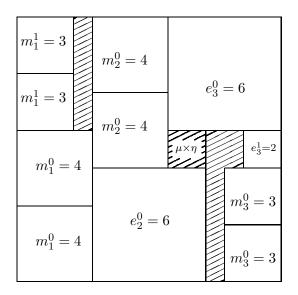


FIGURE 1. Overlapping  $e_2$  and  $e_3$  boxes. The box dimension (leftover shaded part) is 20, the linear correction is  $\ell = r - e_2^0 - e_3^0 + e_3^1 = 4$  so the dimension in the orthogonal case is (20-4)/2 = 8 and in the symplectic case it is (20+4)/2 = 12.

*Proof.* For each i the the  $e_i^j$  are in decreasing order so

$$0 \le \sum_{i} (-1)^{j} e_{i}^{j} \le e_{i}^{0},$$

and similarly for  $f_i^j$ . We have  $e_1^0+e_2^0+e_3^0\leq r$  by Lemma 4.2, and  $e_i^0$  is defined to be bigger than  $f_i^0$  so  $f_1^0+f_2^0+f_3^0\leq r$ . This gives

$$0 \le \sum_{i,j} (-1)^j (e_i^j + f_i^j) \le 2r$$

so 
$$|\ell| \leq r$$
.

The following fact will be used often below. Recall that the quantities being manipulated correspond to twice the dimension.

Corollary 5.3. If the relative character variety has dimension d then  $\beta \leq r + 2d$ .

5.3. **A lemma.** We consider in this section partitions  $\lambda_1 + \ldots + \lambda_k = r$  provided with the decoration of a sequence  $\varepsilon^1, \ldots, \varepsilon^k$  with  $\varepsilon^j \in \{-1, 0, 1\}$ . The following lemma formulates in a general way the computation that leads to bounds in the rectangles of size  $q_i \times r_i$ . The letters here are different from everywhere else.

**Lemma 5.4.** Given N > 0 there exist Q > 0 and P > 0 such that the following property holds. Suppose q, r are integers with  $q \geq Q$  and  $r \leq Nq$ , and suppose given  $\varepsilon = \pm 1$ . Suppose given a partition  $r = \lambda_1 + \ldots + \lambda_k$ ,  $1 \leq \lambda_i \leq q$ , with decoration  $\epsilon_1, \ldots, \epsilon_k$  such that  $\epsilon_i \in \{-1, 0, 1\}$ . Define

$$\Delta := \sum_{i=1}^{k} \lambda_i (q + \epsilon_i - \lambda_i) - \varepsilon q.$$

Then either  $\Delta \geq Nq$  or else there are integers n and p such that  $\Delta = nq - p$  with  $0 \leq p \leq P$ . In the second case

$$\sum_{i=1}^{k} \min(\lambda_i, q - \lambda_i) \le P.$$

*Proof.* Set  $\xi_i := \min(\lambda_i, q + \epsilon_i - \lambda_i)$ , so

$$\lambda_i(q + \epsilon_i - \lambda_i) = \xi_i(q + \epsilon_i - \xi_i),$$

thus

$$\Delta = \sum_{i=1}^{k} \xi_i (q + \epsilon_i - \xi_i) - \varepsilon q.$$

Assuming Q big enough, we have  $q + \epsilon_i - \xi_i \ge (q-1)/2 \ge q/3$ .

We have  $\xi_i \geq -1$  and if  $\xi_i = -1$  then  $q + \epsilon_i - \xi_i = \lambda_i = q$ . Write  $\Delta = \Delta^- + \Delta^+$  with

$$\Delta^{-} = \sum_{\xi_i = -1} \xi_i (q + \epsilon_i - \xi_i) - \varepsilon q$$

and

$$\Delta^{+} = \sum_{\xi_i > 1} \xi_i (q + \epsilon_i - \xi_i).$$

The terms where  $\xi_i = 0$  do not contribute so they are not included here.

As  $\Delta^-$  is a sum of terms equal to  $\pm q$ , there is an integer  $n^-$  such that  $\Delta^- = n^- q$ . Note also that if  $\xi_i = -1$  then it has to be the case  $\lambda_i = q$  and  $\epsilon_i = -1$ . The partition relation  $r = \sum_j \lambda_i$  implies that there are  $\leq r/q$  such terms, to which the extra term  $-\varepsilon q$  should be added. Therefore  $|\Delta^-| \leq r + q$ .

If  $\Delta \geq Nq$  then it is the first case and the proof is done.

Suppose  $\Delta < Nq$ , then the previous bound implies  $\Delta^+ < (N+1)q + r \le (2N+1)q$ . The estimate  $q + \epsilon_i - \xi_i \ge q/3$  implies that

$$\Delta^+ \ge \left(\sum_{\xi_i \ge 1} \xi_i\right) q/3.$$

Combining with  $\Delta^+ < (2N+1)q$  this gives

$$\sum_{\xi_i > 1} \xi_i \le 3(2N+1).$$

Set  $n^+ := \sum_{\xi_i \ge 1} \xi_i$  and separate into two terms

$$\Delta^{+} = n^{+}q - p = n^{+}q - \sum_{\xi_{i} > 1} \xi_{i}(\xi_{i} - \epsilon_{i}).$$

This gives the number p for the conclusion of the lemma, whereas the number n for the conclusion is  $n = n^- + n^+$  For the terms in p we have  $0 \le (\xi_i - \epsilon_i) \le \xi_i + 1 \le 6N + 4$ , so this gives

$$0 \le p \le (6N+4) \sum_{\xi_i > 1} \xi_i \le (6N+4)^2.$$

For the first part of the lemma, taking any  $P \ge (6N+4)^2$  will therefore suffice.

For the second statement, set  $\rho_i := \min(\lambda_i, q - \lambda_i)$ . By adjusting Q we may assume that 3(2N+1) < q-2-3(2N+1) so  $\rho_i$  is either  $\xi_i$  or  $\xi_i - \epsilon_i$ . If  $\xi_i \ge 1$  then  $\rho_i \le 2\xi_i$  so

$$\sum_{\mathcal{E}_i > 1} \rho_i \le 6(2N + 1).$$

We need to bound the sum of  $\rho_i$  where  $\xi_i = 0, -1$ . Since  $\lambda_i \ge 1$ , in this case  $\xi_i = q + \epsilon_i - \lambda_i$  so  $\lambda_i \ge q - 1$ . There are at most  $r/(q-1) \le 2N$  such terms, and for these terms we have  $\rho_i \le 1$ . Thus

$$\sum_{\xi_i \le 0} \rho_i \le 2N,$$

and combining with the other terms we get

$$\sum_{i=1}^{k} \rho_i \le 7(2N+1).$$

Thus, as soon as  $P \ge 7(2N+1)$  we get the second part of the statement in the case  $\Delta < Nq$ .

**Corollary 5.5.** Suppose given N > 1 and  $\eta > 0$ . Then there exists Q > 0 such that the following holds. Suppose  $q \geq Q$  and  $q \leq r \leq Nq$ , suppose  $r = \lambda_1 + \ldots + \lambda_k$  is a partition with  $1 \leq \lambda_i \leq q$ , and suppose  $\epsilon_i \in \{-1, 0, 1\}$ . Let  $\Delta$  denote the quantity of the lemma. Let B denote the number of big parts, that is to say the number of  $\lambda_i \geq q/2$ . If  $\Delta < Nq$ , then  $B \geq 1$  and

$$\left| \frac{q}{r} - \frac{1}{B} \right| < \eta.$$

*Proof.* Let Q' and P be the numbers given by the lemma. From the hypothesis  $\Delta < Nq$ , we are in the second case of the lemma. Let  $\rho_i := \min(\lambda_i, q - \lambda_i)$ . Decomposing into small and big parts, we have

$$r = \sum_{i} \lambda_{i} = \sum_{\text{small}} \lambda_{i} - \sum_{\text{big}} (q - \lambda_{i}) + Bq = \sum_{\text{small}} \rho_{i} - \sum_{\text{big}} \rho_{i} + Bq.$$

Thus

$$|r - Bq| \le \sum_{i} \rho_i \le P.$$

If Q > P, noticing the hypothesis  $r \ge q \ge Q$ , we get  $B \ge 1$ . Divide the above inequality by  $rB \ge Q$  to get

$$\left|\frac{1}{B} - \frac{q}{r}\right| \le \frac{P}{rB} \le \frac{P}{Q}.$$

Choose  $Q > P/\eta$  to get the stated estimate.

#### 6. Non-overlapping case

Assume in this section that we are in non-overlapping case of Proposition 5.1. Hence

$$r = q_1 + q_2 + q_3$$

and the blocks are contained in the corresponding columns. Assume  $q_1 \ge q_2 \ge q_3$ .

There are several different closely related quantities that will be needed. For the *i*-th column, the various parts  $e_i^u$ ,  $f_i^u$  and twice each  $m_i^u$  combine together to give a partition denoted

$$r = \lambda_i^1 + \lambda_i^2 + \dots$$

The signs occurring in the dimension formula yield a decoration of this partition by  $\epsilon_i^j \in \{-1, 0, 1\}$ , such that the linear correction is

$$\ell = r - \sum_{i,j} \epsilon_i^j \lambda_i^j.$$

Let  $\rho_i^j = \min(\lambda_i^j, q_i - \lambda_i^j)$ , so the box-dimension is

$$\beta = \sum_{i,j} \rho_i^j (q_i - \rho_i^j).$$

Recall that the full dimension (multiplied by 2) is

$$\Delta = \beta - \varepsilon \ell = \sum_{i,j} \rho_i^j (q_i - \rho_i^j) - \varepsilon r + \varepsilon \sum_{i,j} \epsilon_i^j \lambda_i^j.$$

Decompose this as  $\Delta = \sum_i \Delta_i$  with

$$\Delta_i = \sum_j \rho_i^j (q_i - \rho_i^j) - \varepsilon q_i + \varepsilon \sum_j \epsilon_i^j \lambda_i^j$$

$$= \sum_{j} \lambda_{i}^{j} (q_{i} - \lambda_{i}^{j}) - \varepsilon q_{i} + \varepsilon \sum_{j} \epsilon_{i}^{j} \lambda_{i}^{j}.$$

The terms in the second expression may then be combined as

$$\Delta_i = \sum_j \lambda_i^j (q_i + \varepsilon \epsilon_i^j - \lambda_i^j) - \varepsilon q_i.$$

Let  $\xi_i^j := \min(\lambda_i^j, q_i + \varepsilon \epsilon_i^j - \lambda_i^j)$ . These are the quantities used in Lemma 5.4 for the *i*-th column of width  $q_i$ , along with

$$\Delta_i = \sum_{j} \xi_i^j (q_i + \varepsilon \epsilon_i^j - \xi_i^j) - \varepsilon q_i.$$

We say that a part is big if  $\lambda_i^j \geq q/2$  and small otherwise. Let  $B_i$  be the number of big parts in the partition  $(\lambda_i^j)_j$ . If  $B_i \geq 1$  define

$$\alpha_i := r - B_i q_i.$$

Lemma 6.1. This satisfies the bound

$$|\alpha_i| \le \sum_j \rho_i^j$$
.

*Proof.* If j is a big part in the ith column then  $\lambda_i^j = q_i - \rho_i^j$ , whereas  $\lambda_i^j = \rho_i^j$  otherwise. Decompose the partition relation into the sum over big and small parts

$$r = \sum_{j} \lambda_{i}^{j} = \sum_{\text{small}} \rho_{i}^{j} + \sum_{\text{big}} (q_{i} - \rho_{i}^{j})$$
$$= B_{i}q_{i} + \sum_{\text{small}} \rho_{i}^{j} - \sum_{\text{big}} \rho_{i}^{j}.$$

Recall that  $\rho_i^j$  are positive, so  $|r - B_i q_i| \leq \sum \rho_i^j$ .

Assume that R is large enough with respect to d, in particular R > 2d. As in Lemma 5.2,  $\Delta_i \ge -q_i$  because of the alternating nature of the signs in the decreasing block sizes of the series  $e_i^j$  and  $f_i^j$ . Therefore  $\Delta_i = \Delta - \sum_{j \ne i} \Delta_j \le 2d + r < 2r$ .

Apply Lemma 5.4 and Corollary 5.5 to this situation. Notice that the input data and the quantity  $\Delta_i$  are the same as there. By the order assumption,  $q_1 \geq r/3$ . Taking  $N \geq 6$ , we find that if  $\Delta_1 \geq Nq_1$ , then  $\Delta_1 \geq 2r$  contradicting the previous paragraph. Thus, we are in the second case for the lemma. Taking a small  $\eta > 0$  and choosing Q as in the lemma and the corollary, assume that  $r \geq 3Q$  so  $q_1 \geq Q$ . The number  $B_1$  of big parts at i = 1 is  $\geq 1$  and we have

$$\left|\frac{q_1}{r} - \frac{1}{B_1}\right| < \eta.$$

**Proposition 6.2.** Let  $I \subset \{1, 2, 3\}$  be a subset of indices such that we know  $r \leq 12q_i$  for  $i \in I$ . Apply Lemma 5.4 with N = 24 and let P, Q be the resulting bounds. For  $r \geq R$  with R big enough with respect to d, P, Q, we have  $B_i \geq 1$  for  $i \in I$ , and

$$\sum_{i \in I} \sum_{j} (\rho_i^j / B_i) \le 1.$$

Therefore  $\sum_{i \in I} |\alpha_i|/B_i \le 1$ .

*Proof.* By hypothesis  $q_i \geq r/12$ , and as we have seen above  $\Delta_i < 2r$ . Thus  $\Delta_i < 24q_i$ , so we are in the second case in the application of Lemma 5.4 to the *i*th column (for each  $i \in I$ ) with N = 24. We get the bounds given there with P and Q, for example

$$\sum_{i} \rho_i^j \le P.$$

In particular, as soon as R > P it cannot be the case that all the parts are small, so  $B_i \ge 1$ . Also note that  $B_i \le 24$ .

By Lemma 6.1, 
$$|(r/B_i) - q_i| \le P/B_i \le P$$
. Also  $|(q_i - \rho_i^j) - q_i| \le P$ , so  $(q_i - \rho_i^j) \ge (r/B_i) - 2P \ge (1 - (2PB_i/R))r/B_i$ .

Thus

$$\sum_{i \in I} \sum_{j} \rho_i^j / B_i \le (1 - (2PB_i/R))^{-1} r^{-1} \sum_{i \in I} \rho_i^j (q_i - \rho_i^j) \le (1 - (2PB_i/R))^{-1} r^{-1} (r + 2d).$$

If R is big with respect to d and  $2PB_i \leq 48P$  (the latter being because  $B_i \leq 24$ ), then for  $r \geq R$  we have  $(1 - (2PB_i/R))^{-1}r^{-1}(r+2d) < 1 + \eta$  for  $\eta$  small. It follows that

$$\sum_{i \in I} \sum_{j} \rho_i^j / B_i \le 1.$$

The bound that is needed for  $\eta$  depends on the lcm of the  $B_i$  which is bounded by something big but fixed given the condition  $B_i \leq 24$ , however in practice the lcm in question is  $\leq 6$ .

The last statement follows by Lemma 6.1.

6.1. The case  $B_1 = 1$ . Assume first that  $B_1 = 1$ . Applying Proposition 6.2 with  $I = \{1\}$  gives  $|\alpha_1| \le 1$  so  $q_1 \in \{r-1, r, r+1\}$ . These are all impossible due to the relation  $r = q_1 + q_2 + q_3$  with  $q_i \ge 1$ . This shows that  $B_1 \ge 2$ .

6.2. The case  $B_1=2$  and  $B_2=2$ . Assume  $B_1=2$ . Applying the proposition at  $I=\{1\}$  gives  $|\alpha_1|\leq 2$ , and  $q_2+q_3=r-q_1=(r+\alpha_1)/2$ . From the order assumption  $q_2\geq q_3$  this gives  $q_2\geq r/5$  so the proposition may be applied at  $I=\{1,2\}$ , yielding  $B_2\geq 1$ . But the case  $B_2=1$  is not possible for the same reason as in the previous paragraph. Thus  $B_2\geq 2$ . Assume for the moment that  $B_2=2$ . Then by Proposition 6.2,  $|\alpha_1|+|\alpha_2|\leq 2$  so

$$2q_3 = 2r - 2q_1 - 2q_2 = (r - 2q_1) + (r - 2q_2) = \alpha_1 + \alpha_2 \le 2.$$

As  $q_3 \ge 1$  this implies that in fact  $q_3 = 1$  and equality holds namely  $\alpha_1 + \alpha_2 = 2$  and  $\alpha_1, \alpha_2 \ge 0$ . The order relation  $q_1 \ge q_2$  implies  $\alpha_1 \le \alpha_2$  so there are two possibilities for  $(\alpha_1, \alpha_2)$ , either (0, 2) or (1, 1). Denote  $q := q_2$ .

All the parts over i=3 have to have size  $\lambda_3^i=1$  so the box dimension is 0 at i=3. The linear correction from this column has absolute value  $\leq 2$  since it is a sum of at most two alternating series of  $\pm 1$ 's.

Proposition 6.2 also gives

$$\sum_{i=1,2} \sum_j \rho_i^j \le 2.$$

Equality must hold here and in the bound of Lemma 6.1, and looking at the proof of that lemma, all the parts with  $\rho_i^j \neq 0$  must be small. In other words, the big parts have size  $q_i$ .

Suppose  $\alpha_1 = 0$  and  $\alpha_2 = 2$ . Then r = 2q + 2 and  $q_1 = q + 1$ . From the previous paragraph it follows that the partition at i = 1 is  $(\lambda_i^1, \lambda_i^2) = (q + 1, q + 1)$  and the partition at i = 2 is either (q, q, 1, 1) or (q, q, 2). The box dimension at i = 1 is 0 and at i = 2 it is 2q - 2 or 2q - 4 respectively.

The linear correction needs to cancel out the terms of size 2q so it has to be approximately -r. Assuming  $r \geq R$  is big enough, this can only happen if the big blocks of size q are not included in the linear correction, or at least they occur with alternating signs. And it has to be the case that the correction starts with  $-\varepsilon r = -r$  so  $\varepsilon = 1$ , i.e. it is the orthogonal case.

Now the remaining part of the linear correction at i=2 consists of a sum of at most two alternating series among the small parts, hence it is at most 2 in absolute value. The absolute value of the term  $\sum_{i,j} (-1)^u (e_i^u + f_i^u)$  is therefore bounded by 4. The dimension is the sum of terms, the first being 2q-2 or 2q-4, the second being -2q-2, and the third being of absolute value  $\leq 4$ . The result is  $\leq 0$ . This does yield a candidate possibility for the rigid case d=0, however it can not lead to a case of dimension  $d\geq 2$ . This rules out  $(\alpha_1,\alpha_2)=(0,2)$ .

Suppose  $\alpha_1 = \alpha_2 = 1$ . Then r = 2q + 1 and  $q_1 = q_2 = q$ . The partitions are both of the form (q,q,1). The box dimension is q-1 in each column for a total of 2q-2. As in the previous case, the linear correction has to start with  $-\varepsilon r = -r$  so  $\varepsilon = 1$  and the big blocks contribute 0 to the linear correction (either because they are paired eigenvalues or because they occur in an alternating series). The small blocks have to contribute values of 1 to the linear correction at i=1,2. The contribution at i=3 is nonzero since there are an odd number of blocks, and the sum of the number of elements in the two alternating series has to be odd. Thus, one series is odd and the other is even, giving a value of 1 for the linear correction over i=3 too. The dimension is

$$2q - 2 - r + 1 + 1 + 1 = 0.$$

This is again a candidate for a rigid solution but does not contribute to any case of  $d \ge 2$ . This completes the proof that the case  $B_1 = B_2 = 2$  doesn't contribute.

6.3. The remaining cases. The two previous subsections show that  $B_1 \ge 2$  and  $B_2 \ge 3$ . Applying Lemmas 5.4 and 6.1 yields  $|\alpha_1| + |\alpha_2| \le P$  for P fixed with respect to R. Thus

$$q_3 = r - q_1 - q_2 = r - \frac{r - \alpha_1}{B_1} - \frac{r - \alpha_2}{B_2} \ge \left(1 - \frac{1}{B_1} - \frac{1}{B_2}\right)r - P.$$

For  $r \geq R$  large we can say  $q_3 \geq r/7$  and apply Proposition 6.2. This says  $B_3 \geq 1$ , and

$$\sum_{i} |\alpha_i|/B_i \le \sum_{i,j} \rho_i^j/B_i \le 1.$$

Now

$$r = q_1 + q_2 + q_3 = \frac{r - \alpha_1}{B_1} + \frac{r - \alpha_2}{B_2} + \frac{r - \alpha_3}{B_3}$$

so

$$\left| \frac{1}{B_1} + \frac{1}{B_2} + \frac{1}{B_3} - 1 \right| \le \sum_i |\alpha_i| / rB_i \le 1/r$$

hence

$$\frac{1}{B_1} + \frac{1}{B_2} + \frac{1}{B_3} = 1.$$

The triple  $(B_1, B_2, B_3)$  is one of (3, 3, 3), (2, 2, 4) or (2, 3, 6), noting that the order relation on  $q_i$  implies  $B_1 \leq B_2 \leq B_3$ .

Recall that we know  $\Delta_i < 2r$  for each column, and  $q_i \geq r/7$  so apply Lemma 5.4 with N=14. Choose the appropriate Q and suppose that  $r \geq 7Q$ , so  $q_i \geq Q$ . The second case of this lemma holds for all three columns.

The lemma says that the dimension is

$$\Delta = \Delta_1 + \Delta_2 + \Delta_3 = n_1 q_1 + n_2 q_2 + n_3 q_3 - p_1 - p_2 - p_3$$

with  $n_i$  integers and  $0 \le p_i \le P$ . The conditions  $\Delta_i \ge -q_i$  that come from the alternating nature of the  $\epsilon_i^j$  imply that  $n_i \ge -1$ . Also  $n \le 15$ .

Our condition is  $\Delta = 2d$  with  $d \ge 2$  fixed with respect to the choice of r. Notice that if Q is large with respect to P, we get a relation of the form

$$\frac{n_1}{B_1} + \frac{n_2}{B_2} + \frac{n_3}{B_3} = 0.$$

Recall that  $(B_1, B_2, B_3)$  is in one of the following cases: (3,3,3), (2,2,4) or (2,3,6). In the (2,3,6) case this gives  $3n_1+2n_2+n_3=0$ , so  $n_3=-3n_1-2n_2$ . With  $n_i\geq -1$  that gives  $n_3\leq 5$ . The bounds for  $n_1$  and  $n_2$  are smaller, as are the bounds for the  $n_i$  in the other two cases. In all, this shows that  $-1\leq n_i\leq 5$ .

Set  $\varphi_i := -\alpha_i/B_i$  so the bound is  $|\varphi_1| + |\varphi_2| + |\varphi_3| \le 1$  and  $q_i = (r/B_i) + \varphi_i$ . The relations  $r = q_1 + q_2 + q_3$  and  $\sum (1/B_i) = 1$  yield

$$\varphi_1 + \varphi_2 + \varphi_3 = 0.$$

The formula of Lemma 5.4 for the dimension becomes

$$\Delta = \frac{n_1 r}{B_1} + \frac{n_2 r}{B_2} + \frac{n_3 r}{B_3} + n_1 \varphi_1 + n_2 \varphi_2 + n_3 \varphi_3 - p_1 - p_2 - p_3.$$

The relation between  $n_i$  and  $B_i$  makes the first terms vanish so

$$\Delta = n_1 \varphi_1 + n_2 \varphi_2 + n_3 \varphi_3 - p_1 - p_2 - p_3.$$

From this, subtract twice the relation  $\varphi_1 + \varphi_2 + \varphi_3 = 0$  to get

$$\Delta = (n_1 - 2)\varphi_1 + (n_2 - 2)\varphi_2 + (n_3 - 2)\varphi_2 - p_1 - p_2 - p_3.$$

Now  $|n_i - 2| \leq 3$  so

$$\left| \sum_{i} (n_i - 2)\varphi_i \right| \le 3 \sum_{i} |\varphi_i| \le 3.$$

The  $p_i$  are positive, so  $\Delta \leq 3$  under the hypothesis  $r \geq R$ . But this contradicts the condition  $\Delta \geq 4$  (recall  $\Delta = 2d$  with d > 0 and d even by Remark 4.8). This proves that the only examples occur for  $r \leq R$ , completing the proof of Theorem 4.9 and hence Theorem 1.1 in the non-overlapping case.

#### 7. The overlapping case

The overlapping case is, up to permuting i, when  $m_1^0 \geq e_1^0$ ,  $e_2^0 \geq m_2^0$ ,  $e_3^0 \ge m_3^0$ , and

$$m_1^0 + e_2^0 + e_3^0 > r$$
.

In particular both  $e_2^0$  and  $e_3^0$  are > 0. Let  $\mu := m_1^0 + e_2^0 + e_3^0 - r$  be the overlapping amount. We note that  $m_1^0 \le \mu$ . Furthermore  $e_2^0 + e_3^0 \le r$ , indeed if it were strictly greater than rthen we would have a reducible system, by looking at the sum of ranks of the monodromy matrices at  $t_2$  and  $t_3$ .

Let  $\nu := r - e_2^0 - e_3^0$ . We note that the box dimension includes a piece of size  $\mu \times \nu$  between the boxes. Define  $q_1 := m_1^0$ ,

$$q_2 := e_2^0 - \mu = r - m_1^0 - e_3^0, \quad q_3 := e_3^0 - \mu = r - m_1^0 - e_2^0.$$

Also set  $r_1 := r$  and

$$r_2 := e_3^0 + \nu = r - e_2^0 = q_1 + q_3, \quad r_3 := e_2^0 + \nu = r - e_3^0 = q_1 + q_2.$$

Order the points  $t_2, t_3$  so that  $q_2 \geq q_3$ .

By Proposition 5.1, the squares that remain after  $e_2^0 \times e_2^0$  and  $e_3^0 \times e_3^0$  fit into the three rectangles of sizes  $q_1 \times r$ ,  $q_2 \times r_2$  and  $q_3 \times r_3$ . In particular  $q_i \geq m_i^j$ ,  $q_i \geq f_i^j$ , and  $q_i \geq e_i^u$  for  $u \geq 1$  (as well as for u = 0 if i = 1).

The leftover region whose area is the box dimension, consists of the rectangle of size  $\mu \times \nu$ , plus the leftover pieces in those three rectangles.

It is suggested to look at Figure 1 that has a picture of a typical such arrangement. In that picture we have  $r=14, q_1=q_2=q_3=4, \mu=\nu=2$ and  $r_2 = r_3 = 8$ . The leftover region in the  $q_1 \times r$  rectangle has area 6, there is no leftover region in the  $q_2 \times r_2$  rectangle, and the leftover region in the  $q_3 \times r_3$  rectangle has area 10. To these we add  $\mu\nu = 4$  so the total box dimension is  $\beta = 20$ .

**Lemma 7.1.** The following relations hold:

- $q_1 = \mu + \nu$ , thus  $\mu, \nu \le q_1$ ,  $q_2 + \mu = e_2^0$  and  $q_3 + \mu = e_3^0$ ,
- $q_1 + q_2 + q_3 + \mu = r$  or equivalently  $q_2 + q_3 + 2\mu + \nu = r$ ,
- $\bullet \ e_2^{\bar{0}} + e_3^{\bar{0}} + \bar{\nu} = r,$
- $e_1^j, f_1^j < \nu$ ,

•  $\mu \leq q_3 \leq q_2$ .

Proof. The equalities follow from the definitions. For  $q_i \geq 1$  we note that the partitions at each point need to have at least 2 parts, and since the remaining ones after  $e_2^0$  and  $e_3^0$  fit into the rectangles, we can not have any  $q_i = 0$ . The conditions  $e_1^j$ ,  $f_1^j \leq \nu$  come from the bound of Lemma 4.2. For the last part recall the order relation saying  $q_2 \geq q_3$ . Lemma 4.4 gives the cruder estimate  $m_1^0 + e_2^0 + (e_3^0/2) \leq r$ , and since  $q_1 = m_1^0$  and  $e_i^0 = q_i + \mu$  for i = 2, 3 we have  $q_1 + q_2 + \mu + (q_3 + \mu)/2 \leq q_1 + q_2 + q_3 + \mu$  so  $\mu \leq q_3$ .

7.1. A numerical modification. We define a modification of the partition data that preserves the dimension and leads into the non-overlapping case. This looks like something that could come from a middle convolution operation involving distinguished blocks  $m_1^0$ ,  $e_2^0$  and  $e_3^0$ , see Question 8.3 below.

The rank changes from r to  $\tilde{r} := r - \mu$ .

Given partitions  $(\lambda_i^j)_j$  with decorations  $\epsilon_i^j$ , suppose  $\lambda_1^1 = m_1^0$  and  $\lambda_i^1 = e_i^0$  for i=2,3. Then define new partitions  $(\tilde{\lambda}_i^j)_j$  and new decorations  $\tilde{\varepsilon}_i^j$  as follows. If  $j \neq 1$  put  $\tilde{\lambda}_i^j := \lambda_i^j$  and at j=1 set  $\tilde{\lambda}_i^j := \lambda_i^j - \mu$ . For any j at i=2,3, set  $\tilde{\varepsilon}_i^j := \epsilon_i^j$ . At i=1 suppose that  $\lambda_1^2 = m_1^0$  is the other part in the conjugate pair. Recall that  $\epsilon_1^1 = \epsilon_1^2 = 0$  since these parts correspond to paired eigenvalues. For  $j \neq 1,2$  set  $\tilde{\varepsilon}_j^i := \epsilon_j^j$ . Then put

$$\epsilon_1^2:=1, \quad \epsilon_1^1:=-1.$$

In more descriptive terms, at i=2,3 we replace the part  $e^0_i$  by  $\tilde{e}^0_i:=e^0_i-\mu$ , and keep the other parts whose blocks go in the rectangles of size  $q_i\times r_i$ , the same. Over i=1 the pair  $(m^0_1,m^0_1)$  is changed into  $\tilde{e}^0_1:=m^0_1$  and  $\tilde{e}^1_1:=m^0_1-\mu=\nu$ . The other parts in the series  $e^j_1$  are then added on later:  $\tilde{e}^j_1:=e^{j-2}_1$ . For the second series  $\tilde{f}^j_i:=f^j_i$ . The picture of the result of applying this modification to the example of Figure 1 is shown in Figure 2.

Notice that the inequalities  $e_1^j \leq \nu$  imply that the new  $\tilde{e}_1^j$  remain in decreasing order.

**Lemma 7.2.** The box dimension and linear correction associated to the new decorated partition data  $(\tilde{\lambda}, \tilde{\epsilon})$  are equal to those associated to  $(\lambda, \epsilon)$ . Hence, the dimension associated to the new data is the same as the dimension associated to the previous ones. The new data fit into the non-overlapping case, and we have  $\tilde{r} \geq 2r/3$ .

*Proof.* The new blocks fit into three columns of width  $q_1, q_2, q_3$  and notice that  $q_1 + q_2 + q_3 = \tilde{r}$ . For i = 2, 3 the block dimension in the rectangle  $q_i \times \tilde{r}$  is the same as that of the rectangle  $q_i \times r_i$  since  $\tilde{r} = r_i + \tilde{e}_i^0$  and the first block of size  $\tilde{e}_i^0 = q_i$  does not leave any leftover area. In the rectangle of size  $q_1 \times \tilde{r}$  the leftover area consists of the same configuration of leftover area as in the rectangle of size  $q_1 \times r$  (which is contained in fact in a rectangle of

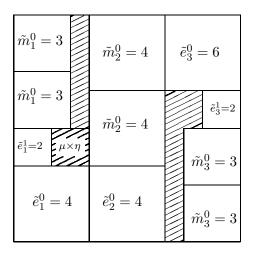


FIGURE 2. The picture after modifying the example of Figure 1.

size  $q_1 \times (r - 2m_1^0)$ ) plus a rectangle of size  $\mu \times \nu$  that is next to the block of size  $\nu \times \nu$  as  $\nu = \tilde{e}_1^1$ .

The previous linear correction was

$$r - \sum_{i,j} (-1)^j (e_i^j + f_i^j).$$

For the new linear correction notice that the signs of all the parts that come from the previous one remain the same, but there are two new parts  $\tilde{e}_1^1 = q_1 = m_1^0$  and  $\tilde{e}_1^1 = m_1^0 - \mu = \nu$ . So for the new linear correction,

$$\sum_{j}(-1)^{j}(\tilde{e}_{1}^{j}+\tilde{f}_{1}^{j})=(m_{1}^{0}-(m_{1}^{0}-\nu))+\sum_{j}(-1)^{j}(e_{1}^{j}+f_{1}^{j})=\mu+\sum_{j}(-1)^{j}(e_{1}^{j}+f_{1}^{j})$$

whereas for i = 2, 3

$$\sum_{j} (-1)^{j} (\tilde{e}_{i}^{j} + \tilde{f}_{i}^{j}) = (e_{i}^{0} - \mu) + \sum_{j \geq 1} (-1)^{j} e_{i}^{j} + \sum_{j} (-1)^{j} f_{i}^{j} = -\mu + \sum_{j} (-1)^{j} (e_{i}^{j} + f_{i}^{j}).$$

Thus

$$\sum_{i,j} (-1)^j (\tilde{e}_i^j + \tilde{f}_i^j) = (\mu - \mu - \mu) + \sum_{i,j} (-1)^j (e_i^j + f_i^j)$$

and we get

$$\tilde{r} - \sum_{i,j} (-1)^j (\tilde{e}_i^j + \tilde{f}_i^j) = (r - \mu) + \mu - \sum_{i,j} (-1)^j (e_i^j + f_i^j)$$

which is equal to the previous linear correction. The (twice) dimension is equal to the box dimension minus  $\varepsilon$  times the linear correction, so this is the same for the new data as for the previous data.

In the new data we have three non-overlapping columns of widths  $q_1, q_2, q_3$  and all  $\tilde{\lambda}_i^j \leq q_i$ . Thus, the new data fits into the non-overlapping case. Finish the proof by noting that  $q_2 \geq q_3 \geq \mu$  and  $r \geq q_2 + q_3 + \mu$  so  $\mu \leq r/3$  and  $\tilde{r} = r - \mu > 2r/3$ .

Corollary 7.3. Let  $R^{\text{non}}(d)$  be the bound of Theorem 4.9 for the non-overlapping case, such that  $r \leq R^{\text{non}}(d)$  for any d-dimensional numerically MC-minimal decorated partition data in the non-overlapping case. Then, for any numerically MC-minimal decorated partition data in the overlapping case we have  $r \leq (3/2)R^{\text{non}}(d)$ .

Proof. The modification  $(r, \lambda, \varepsilon) \mapsto (\tilde{r}, \tilde{\lambda}, \tilde{\varepsilon})$  sends numerically MC-minimal partition data in the overlapping case to numerically MC-minimal partition data in the non-overlapping case, preserving the dimension. Thus  $\tilde{r} \leq R^{\text{non}}(d)$ . As  $\tilde{r} \geq 2r/3$ , this gives  $2r/3 \leq R^{\text{non}}(d)$  which is the desired bound.

This corollary completes the proof of Theorem 4.9 in the overlapping case, hence completing the proof of Theorem 1.1.

A remark is perhaps in order about this method of proof to treat the overlapping case. In general, the reductions by middle convolution implicit in the main reduction lemmas will happen in an iterative sequence. If some inequality is not satisfied, then an MC operation is applied to reduce the rank. Along the way, an impossible situation (i.e. partitions that contradict irreducibility such as in the first parts of the proofs of the lemmas) may occur, and that implies that the original relative character variety was empty. If no such impossible situation happens along the way, eventually the result will be numerically MC-minimal, or else a rank 1 or rank 2 hypergeometric system, since the rank decreases at each step. This is the essence of Katz's algorithm [41].

Before applying the sequence of reductions, the rank could have been arbitrarily big. Indeed, one could apply various quadratic middle convolutions that increase the rank, as was used recently in [45].

In the present section, we used a sort of virtual reduction of the rank that modifies the partition data. This construction does not provide an isomorphism between the character varieties. Fortunately, it doesn't need to be applied iteratively, indeed the result directly becomes numerically MC-minimal and non-overlapping. This gives a bound on the ratio of the ranks before and after the modification.

#### 8. Further questions

**Remark 8.1.** It seems likely that the bound R(d) should be linear in d and explicit, for example probably R(d) = 200d will do. However, we have not verified the details of this.

**Problem 8.2.** Using the bound, provide lists of the possible cases by computer.

There are clearly many other directions for further work. In the first place, to consider an arbitrary number  $k \geq 3$  of singularities. Our main reduction lemmas are tied to the case of k = 3 so more work is needed.

The existence question was discussed in Section 1.1.

Then one should consider other groups. Of course, for the simple exceptional groups there are only finitely many conjugacy classes to check so the finiteness property is automatic. It might nevertheless constitute a difficult computational question to enumerate the possibilities.

On the other hand, for different forms of the classical groups—coverings such as the *Spin* groups or quotients by subgroups of the center—new techniques would be necessary.

One may also fix alternate DM-stack structures over  $\mathbb{P}^1$  with non-trivial generic stabilizers. There are relative character varieties for local systems on such DM-stacks but their combinatorial descriptions will be more complicated.

Question 8.3. Is there a geometric transformation like middle convolution that gives an isomorphism of moduli spaces corresponding to the modification of decorated partition data used in Section 7.1 above?

I did not find a middle convolution operation preserving quadraticity that did that. Recall that we had an operation giving a weaker reduction in rank, in Lemma 4.4. One may wonder if the searched-for operation might involve convolution with an irregular convolution kernel.

There is such an isomorphism in a closely related case. Given partition data in the overlapping case, we can change the two paired eigenvalues of multiplicity  $m_1^0$  into half-unital parts in the partition associated to a unipotent monodromy transformation, putting them at the start of the  $e_1^j$  series. This doesn't modify the dimension, and could probably be done in the parabolic Higgs setting for parabolic weights very close to 0, by forgetting the weights on those pieces. In this case, applying twice the middle convolution operation of Lemma 4.2 furnishes the required modification. On the second convolution the rank does not change (the  $\delta$ -invariant becomes 0) but the second operation puts the arrangement back to the one given by the above modification. Just doing the first operation changes some signs of the eigenvalues  $\pm 1$  leading to a slightly different configuration.

**Dedication:** Gang Xiao was one of the pillars of the Algebraic Geometry group in Nice. We have missed his gentle presence and wry sense of humour. I dedicate this paper to him.

Acknowledgements: I would like to thank Ron Donagi, Chloé Dubas, Vladimir Kostov, Tony Pantev, and Jérémy Toulisse for discussions on topics here. This work was supported by a grant from the Institute for Advanced Study, by the Agence Nationale de la Recherche programs Hodgefun ANR-16-CE40-0011 and 3ia Côte d'Azur ANR-19-P3IA-0002, by the European Research Council Horizons 2020 grant 670624 (Mai Gehrke's DuaLL

project), and by the Simons Collaboration on Homological Mirror Symmetry.

#### References

- [1] Alfaya D, Gómez T. Automorphism group of the moduli space of parabolic bundles over a curve. Advances in Math., 2021, 393: 108070.
- [2] Álvarez-Cónsul L, García-Prada O. Hitchin-Kobayashi correspondence, quivers, and vortices. Comm. Math. Phys., 2003, 238: 1-33.
- [3] Antón-Sancho Á. Fixed points of involutions of G-Higgs bundle moduli spaces over a compact Riemann surface with classical complex structure group. Frontiers of Mathematics, 2024, 1-15.
- [4] Barajas Ayuso G. Fixed points in Higgs bundle moduli spaces and the Prym-Narasimhan-Ramanan construction. PhD dissertation, ICMAT: Universidad Autónoma de Madrid, 2023.
- [5] Balaji V, Seshadri C S. Moduli of parahoric G-torsors on a compact Riemann surface. J. Alg. Geom., 2015, 24-1: 1-49.
- [6] Balasubramanian A, Distler J, Donagi R, Perez-Pardavila C. The Hitchin Image in Type-D. arXiv:2310.05880, 2023.
- [7] Baraglia D, Kamgarpour M, Varma R. Complete integrability of the parahoric Hitchin system. Int. Math. Res. Notices, 2019, 21: 6499-6528.
- [8] Belkale P, Kiers J. Vertices in multiplicative eigenvalue problem for arbitrary groups. arXiv:2306.16676, 2023.
- [9] Bhosle U. Moduli of orthogonal and spin bundles over hyperelliptic curves. Compositio Math., 1984, 51-1: 15-40.
- [10] Bhosle-Desale U. Degenerate symplectic and orthogonal bundles on P. Math. Ann., 1984, 267-3: 347-364.
- [11] Bhosle U, Ramanathan A. Moduli of parabolic G-bundles on curves. Math. Z., 1989, 202-2: 161-180.
- [12] Biquard O, García-Prada O, Mundet i Riera I. Parabolic Higgs bundles and representations of the fundamental group of a punctured surface into a real group. Adv. in Math., 2020, 372: 107305.
- [13] Biswas I, Majumder S, Wong M. Orthogonal and symplectic parabolic bundles. J. Geom. Phys., 2011, 61: 1462–1475
- [14] Boalch P. Riemann-Hilbert for tame complex parahoric connections. Transformation groups, 2011, 16-1: 27-50.
- [15] Boden H, Yokogawa K. Moduli Spaces of parabolic Higgs bundles and parabolic K(D) pairs over smooth curves: I. Int. J. Math., 1996, 07-05: 573-598.
- [16] Bogner M, Reiter S. On symplectically rigid local systems of rank four and Calabi–Yau operators. J. Symb. Computation, 2013, 48: 64-100.
- [17] Bradlow S, García-Prada O, Gothen P. Maximal surface group representations in isometry groups of classical Hermitian symmetric spaces. Geometriae Dedicata, 2006, 122-1: 185-213.
- [18] Bradlow S, García-Prada O, and Mundet i Riera I. Relative Hitchin-Kobayashi correspondences for principal pairs. Quart. J. Math., 2003, 54: 171-208.
- [19] Chakraborty S, Majumder S. Orthogonal and Symplectic Parabolic Connections and Stack of Roots. Bulletin des Sciences Mathématiques, 2024, p. 103397.
- [20] Crawley-Boevey W. Indecomposable parabolic bundles and the existence of matrices in prescribed conjugacy class closures with product equal to the identity. Publ. Math. I.H.E.S., 2004, 100: 171-207.
- [21] Dai S, Li Q. On cyclic Higgs bundles. Math. Ann, 2020, 376-3-4: 1225-1260.
- [22] Dettweiler M, Reiter S. On rigid tuples in linear groups of odd dimension. J. Algebra, 1999, 222-2: 550-560.

- [23] Dettweiler M, Reiter S. An algorithm of Katz and its application to the inverse Galois problem. J. Symbolic Comp., 2000, 30-6: 761-798.
- [24] Dettweiler M, Reiter S. Middle convolution of Fuchsian systems and the construction of rigid differential systems. J. Algebra, 2007, 318-1: 1-24.
- [25] Dettweiler M, Reiter S. On the Hodge theory of the additive middle convolution. Publ. RIMS, 2020, 56-3: 503-537.
- [26] Dettweiler M, Sabbah C. Hodge theory of the middle convolution, Publ. RIMS, 2013, 49-4: 761-800.
- [27] Domokos M. Quiver moduli spaces of a given dimension. Journal of Combinatorial Algebra, 2024, doi: 10.4171/JCA/97.
- [28] Donagi R, Pantev T. Geometric Langlands and non-abelian Hodge theory. Surveys in differential geometry, 2008, 13-1: 85-116.
- [29] Fredrickson L, Mazzeo R, Swoboda J, Weiss H. Asymptotic geometry of the moduli space of parabolic  $SL(2,\mathbb{C})$ -Higgs bundles. J. London Math. Soc., 2022, 106-2: 590-661
- [30] Fujiki A. Hyperkähler structure on the moduli space of flat bundles. In: Prospects in Complex Geometry: Proceedings of the 25th Taniguchi International Symposium held in Katata, and the Conference held in Kyoto, July 31–August 9, 1989. Springer Berlin Heidelberg, 2006, 1-83.
- [31] García-Prada O, Logares M, Muñoz V. Moduli spaces of parabolic U(p,q)-Higgs bundles. Quarterly J. Math., 2009, 60-2: 183-233.
- [32] Gongopadhyay K. On the conjugacy classes in the orthogonal and symplectic groups over algebraically closed fields. Expositiones Mathematicae, 2010, 28-4: 351-356.
- [33] Guralnick R, Larsen M, Tiep P. Character levels and character bounds for finite classical groups. Inventiones Math., 2024, 235: 151-210.
- [34] Hameister T, Morrissey B. The Hitchin fibration for symmetric pairs. arXiv:2407.00961, 2024.
- [35] Hameister T, Ngô B-C. The companion section for classical groups. arXiv:2407.00962, 2024.
- [36] Fernandez Herrero A. On the quasicompactness of the moduli stack of logarithmic G-connections over a curve. arXiv 2002.11832, 2020.
- [37] Hitchin N. Stable bundles and integrable systems. Duke Math. J., 1987, 54-1: 91-114.
- [38] Hitchin N. The self-duality equations on a Riemann surface. Proc. London Math. Soc., 1987, 3-1: 59-126.
- [39] Holdt M. Hyperkähler metrics on the moduli space of weakly parabolic Higgs bundles. arXiv:2106.16017, 2021.
- [40] Konno H. Construction of the moduli space of stable parabolic Higgs bundles on a Riemann surface. J. Math. Soc. of Japan, 1993, 45-2: 253-276.
- [41] Katz N. Rigid local systems. Annals of Math. Studies 139, Princeton University Press, 1996.
- [42] Kostov V. On the Deligne-Simpson problem. Proc. Steklov Inst. Math., 2002, 238: 148-185.
- [43] Kraft H, Procesi C. On the geometry of conjugacy classes in classical groups. Comment. Math. Helvetici, 1982, 57: 539-602.
- [44] Kydonakis G, Sun H, and Zhao L. Tame parahoric Higgs torsors for a complex reductive group. arXiv:2107.01977, 2021.
- [45] Lam Y H J, Landesman A, Litt D. Finite braid group orbits on  $SL_2$ -character varieties. arXiv:2308.01376, 2023.
- [46] Laumon G. Un analogue global du cône nilpotent. Duke Math. J., 1988, 56-1: 647-671.
- [47] Lubotzky A, Majid S. Varieties of representations of finitely generated groups. Memoirs AMS 336, Providence, 1985.
- [48] Martin N. Behaviour of some Hodge invariants by middle convolution. Bulletin de la société mathématique de France, 2021, 149-3: 479-500.

- [49] Martin N. Middle multiplicative convolution and hypergeometric equations. arXiv:1809.08867, 2018.
- [50] Maruyama M, Yokogawa K. Moduli of parabolic stable sheaves. Math. Ann., 1992, 293-1: 77-99.
- [51] Mehta V, Seshadri C S. Moduli of vector bundles on curves with parabolic structures. Math. Ann., 1980, 248-3: 205-239.
- [52] Nakajima H. Hyperkähler structures on moduli spaces of parabolic Higgs bundles on Riemann surfaces. In: Moduli of vector bundles (Sanda, 1994; Kyoto, 1994). Lecture Notes in Pure and Appl. Math. 179, Dekker, New York, 1996, 199-208.
- [53] Ono T. Harmonic bundles with symplectic structures. arXiv:2403.15719, 2024.
- [54] Pandey Y. Criteria for existence of stable parahoric  $SO_n$   $Sp_n$  and Spin bundles on  $\mathbb{P}^1$ . J. Ramanujan Math. Soc., 2014, 29-1: 93-117.
- [55] Reisert P. Moduli spaces of parabolic twisted generalized Higgs bundles. PhD dissertation, LMU München, 2016.
- [56] Roy S. Hitchin fibration on moduli of symplectic and orthogonal parabolic Higgs bundles. Math. Physics, Analysis and Geometry, 2020, 23-4: 41, 13pp.
- [57] Sheng M, Sun H, Wang J. On the gr-semistable filtration of orthogonal/symplectic  $\lambda$ -connections. arXiv:2401.09956, 2024.
- [58] Simpson C. Katz's middle convolution algorithm. Pure Appl. Math. Quarterly, 2009, 5-2: 781-852.
- [59] Simpson C. Local Systems on Proper Algebraic V-manifolds. Pure Appl. Math. Quarterly, 2011, 7-4: 1675-1760.
- [60] Soibelman A. Parabolic bundles over the projective line and the Deligne–Simpson problems. Quarterly J. Math., 2016, 67-1: 75-108.
- [61] Springer T, Steinberg R. Coniugacy classes. Seminar on algebraic groups and related finite groups. Springer Lecture Notes 131, Berlin-Heidelberg-New York, 1970.
- [62] Tholozan N, Toulisse J. Compact connected components in relative character varieties of punctured spheres. Épijournal Géom. Alg., 2021, 5.
- [63] Topley L. Centralisers in Classical Lie Algebras. PhD dissertation, University of Manchester, arXiv:1310.2785, 2013.
- [64] Wang B, Wen X, Wen Y. Springer correspondence and mirror symmetries for parabolic Hitchin systems. arXiv:2403.07552, 2024.
- [65] Wen X. Parabolic Higgs bundles on projective line, quiver varieties and Deligne-Simpson problem. arXiv:2101.01913, 2021.
- [66] Xiao G. On abelian automorphism group of a surface of general type. Inventiones Math., 1990, 102-1: 619-631.
- [67] Xiao G. Bound of Automorphisms of Surfaces of General Type, I Annals of Math., 1994, 139-1: 51-77.
- [68] Yokogawa K. Moduli of parabolic Higgs sheaves. In: Moduli of vector bundles (Sanda, 1994; Kyoto, 1994), Lecture Notes in Pure and Appl. Math. 179, Dekker, New York, 1996,, 287-296.

Email address: carlos.simpson@univ-cotedazur.fr

Université Côte d'Azur, CNRS, LJAD, 06108 Nice Cedex 02, France