Cartesian convexity as the key notion in the variational existence theory for nonlocal supremal functionals

Carolin Kreisbeck [a], Antonella Ritorto [a], Elvira Zappale [b],*

[a] *Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt-Ingolstadt, 85071 Eichstätt, Germany*
[b] *Dipartimento di Scienze di Base ed Applicate per l'Ingengeria/ Sapienza - Università di Roma, 00161, Rome, Italy*

ABSTRACT

Motivated by the direct method in the calculus of variations in L^∞, our main result identifies the notion of convexity characterizing the weakly* lower semicontinuity of nonlocal supremal functionals: Cartesian level convexity. This new concept coincides with separate level convexity in the one-dimensional setting and is strictly weaker for higher dimensions. We discuss relaxation in the vectorial case, showing that the relaxed functional will not generally maintain the supremal form. Apart from illustrating this fact with examples of multi-well type, we present precise criteria for structure-preservation. When the structure is preserved, a representation formula is given in terms of the Cartesian level convex envelope of the (diagonalized) original supremand. This work does not only complete the picture of the analysis initiated in Kreisbeck and Zappale (2020), but also establishes a connection with double integrals. We relate the two classes of functionals via an L^p-approximation in the sense of Γ-convergence for diverging integrability exponents. The proofs exploit recent results on nonlocal inclusions and their asymptotic behavior, and use tools from Young measure theory and convex analysis.

©2022 Elsevier Ltd. All rights reserved.

1. Introduction

The existence theory for minimizers in the calculus of variations is closely linked to convexity notions. As coercivity is usually achieved in function spaces endowed with weak(*) topologies, the central task to guarantee the applicability of the direct method lies in verifying the (sequential) weak(*) lower semicontinuity of the functionals. Over the decades, substantial effort has been put into finding necessary and sufficient conditions for the latter. Depending on the class of functionals at hand, different types of generalized convexity come into play. Fundamental results that identify the correct notions (under suitable growth assumptions) include

---

in the classical integral setting: convexity for integral functionals on Lebesgue spaces [6,17], quasiconvexity for integral functionals on Sobolev spaces with dependence on the weak gradients [2,38], $\mathcal{A}$-quasiconvexity for integral functionals subject to first-order PDE constraints [27], and in the branch of variational L^∞-problems [5,31]: level convexity for supremal functionals on $W^{1,\infty}$ with supremands depending on weak gradients in the scalar case [10,45], and strong Morrey quasiconvexity in the corresponding vectorial case [9].

If functionals fail to be weak($*$) lower semicontinuous, the existence of minimizers is not readily guaranteed. In this case, the study of the relaxed problem, for which one replaces the functional by its weak($*$) lower semicontinuous envelope, allows deducing helpful information about the limits of minimizing sequences. The difficulty is then to find explicit representation formulas for the relaxed functionals that are sufficiently easy to work with. This includes in particular to answer the question whether the formulas are structure-preserving, meaning whether they are again of the same type as the original functionals. The overall rationale for integral and supremal functionals in case structure-preservation holds is then that the relaxation can be obtained, from suitable convexification of its integrand and supremand, respectively; for an introduction to relaxation theory, see e.g. [20,48].

Over the last years, the study of variational models with nonlocal features has attracted increased interest in the community, motivated by the desire to develop a solid understanding of global effects, long-range interactions, and singular behavior in physical phenomena and technical applications, which standard local modeling approaches cannot capture. To mention but a few selected examples from the recent literature, functionals with a nonlocal character appear in the theory of phase transitions [22,49], in peridynamics [13,36], in new models of hyperelasticity [11,12], in image processing [15,23] or in machine learning applications [4,30]. From the mathematical perspective, the presence of nonlocality in variational problems requires substantially different techniques from standard ones, which often rest on localization arguments and are therefore not applicable. Regarding double-integral functionals, essential insights have been established by now. Yet, there remain gaps in the literature, especially when it comes to relaxation theory; we give a brief overview of the current state of the art towards the end of the introduction, see also Table 1. Another type of nonlocal integral functionals, namely, those involving Riesz fractional gradients are studied in [33] (cf. also [12,50]). In this setting, a suitable translation procedure between classical and fractional gradients shows in particular that quasiconvexity of the integrand provides the correct condition for weak lower semicontinuity.

This paper revolves around a class of nonlocal variational problems in L^∞. Precisely, our objects of interest are nonlocal homogeneous supremal functionals of the form

$$J_W(u) = \operatorname*{ess\,sup}_{(x,y)\in \Omega\times\Omega} W(u(x),u(y)), \qquad u \in L^\infty(\Omega;\mathbb{R}^m), \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open set, $m \in \mathbb{N}$ and $W\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is (mostly assumed to be) lower semicontinuous and coercive; observe that the functional J_W is invariant under symmetrization and diagonalization of its supremand W, as first shown in [34, Section 7.1], i.e.,

$$J_W = J_{\widehat{W}} \qquad \text{with } \widehat{W}(\xi,\zeta) = \max W(\{\xi,\zeta\}\times\{\xi,\zeta\}) \text{ for } (\xi,\zeta) \in \mathbb{R}^m \times \mathbb{R}^m;$$

for more details on the role of diagonalization see Section 2.3. As our main contributions, we characterize the L^∞-weak* lower semicontinuity of J_W in terms of a new convexity condition on the symmetrized and diagonalized supremand $\widehat{W}$ and determine criteria for structure preservation under relaxation, see Theorems 1.1 and 1.2 for the precise statements. It follows in particular that relaxations may generally not be supremal functionals anymore in the vectorial setting $m > 1$. Our analysis provides a comprehensive result, unifying and extending the work initiated in [34]. There, two of the authors solve the scalar case ($m = 1$) and consider the vectorial case under rather restrictive assumptions, which, in particular, force the relaxations

to be again of the type (1.1). The arguments in [34] build up around the notion of separate level convexity in the vectorial components. Inspired by the local setting of L^∞-functionals, where one studies functionals

$$L^\infty(\Omega; \mathbb{R}^m) \ni u \mapsto \operatorname*{ess\,sup}_{x \in \Omega} f(u(x))$$

with a coercive and lower semicontinuous $f : \mathbb{R}^m \to \mathbb{R}$, this may appear like a natural choice. Indeed, as already mentioned above, in this local setting, level convexity of the supremand f is known to be a necessary and sufficient condition for weak* lower semicontinuity [10], and relaxation is again a supremal functional with the level convexification of f as the supremand [44], cf. also [1,18].

We show, however, that separate level convexity does not constitute the appropriate concept for characterizing the weak* lower semicontinuity of nonlocal supremals in dimensions greater than one. Addressing the general vectorial case requires a change of perspective along with a new concept of convexity for sets, which we call *Cartesian convexity*:

A set $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is Cartesian convex if $A \times A \subset E$ implies $A^{\mathrm{co}} \times A^{\mathrm{co}} \subset E$;

see Definition 3.1 and Lemma 3.2. Consequently, a function $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is said to be Cartesian level convex, if all its sublevel sets are Cartesian convex; an equivalent representation via a Jensen-type inequality is presented in Proposition 3.11. Note that Cartesian (level) convexity is indeed identical with separate (level) convexity if $m = 1$ and is strictly weaker for higher dimensions, cf. Remark 3.3.

With this terminology at hand, our key characterization result now reads as follows.

Theorem 1.1 (*Characterization of Weak* Lower Semicontinuity*). *Let* $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ *be lower semicontinuous and coercive, and let* J_W *be defined as in* (1.1). *Then,* J_W *is (sequentially) weakly* lower semicontinuous in* $L^\infty(\Omega; \mathbb{R}^m)$ *if and only if*

$$\widehat{W} \text{ is Cartesian level convex,}$$

i.e., all sublevel sets of $\widehat{W}$ *are Cartesian convex.*

Notice that the previous result can equivalently be stated for weak lower semicontinuity in L^∞ and its sequential version, considering that the L^∞-weak* topology admits a metrizable description on bounded sets, cf. e.g. [26, A.1.5]; without further mentioning, we use the same reasoning throughout the paper.

The proof of Theorem 1.1 passes through the theory of nonlocal inclusions. Indeed, by an adaption of the analogous statement for local supremals in [1] (see [34, Proposition 7.1]), the weak* lower semicontinuity of J_W is equivalent to the weak* closedness of all its sublevel sets. The latter are given exactly by the functions $u \in L^\infty(\Omega; \mathbb{R}^m)$ that solve an inclusion problem

$$(u(x), u(y)) \in K \quad \text{for a.e. } (x, y) \in \Omega \times \Omega \tag{1.2}$$

with a compact set $K \subset \mathbb{R}^m \times \mathbb{R}^m$, where K is chosen as the sublevel set $L_c(W)$ of the supremand W for each level $c \in \mathbb{R}$. With this viewpoint, Theorem 1.1 is a consequence of the following statement of independent interest: the set of solutions to (1.2) remains invariant under weak*-limits if and only if the symmetrization and diagonalization of K, i.e.,

$$\widehat{K} =: \{(\xi, \zeta) \in K : \{\xi, \zeta\} \times \{\xi, \zeta\} \in K\},$$

is Cartesian convex, see Proposition 4.1. In showing this, we benefit from recent insights into alternative representations of the solution sets and their asymptotic behavior as established in [34,35]. This allows us

to break the complexity down to the study of nonlocal inclusions that are much easier to handle, namely, problems of the form (1.2) with Cartesian squares of the form $A \times A \subset \mathbb{R}^m \times \mathbb{R}^m$ as constraining sets.

Whenever J_W fails to be weakly* lower semicontinuous in $L^\infty(\Omega; \mathbb{R}^m)$, we pass to the relaxed problem, for which one replaces the functional by its weak* lower semicontinuous envelope; precisely, the relaxation of J_W is given by

$$J_W^{\mathrm{rlx}}(u) := \inf\big\{\liminf_{j\to\infty} J_W(u_j) : u_j \overset{*}{\rightharpoonup} u \text{ in } L^\infty(\Omega; \mathbb{R}^m)\big\} \tag{1.3}$$

for $u \in L^\infty(\Omega; \mathbb{R}^m)$. For the scalar case $m = 1$, it was recently established in [34, Theorem 1.3] that

$$J_W^{\mathrm{rlx}} = J_{\widehat{W}^{\mathrm{slc}}}, \tag{1.4}$$

where the relaxed supremand $\widehat{W}^{\mathrm{slc}}$ is the separate level convexification of $\widehat{W}$. This shows, in particular, that the supremal structure of the functional is invariant under relaxation. The argument in [34] relies on the observation that the separately convex hull of a compact set K (in particular, any sublevel set $L_c(W)$ with $c \in \mathbb{R}$) has a specifically simple structure, namely it can be represented as a union of basic squares with corners in K, in formulas,

$$K^{\mathrm{sc}} = \bigcup_{(\xi,\varsigma)\in K} [\xi,\varsigma] \times [\xi,\varsigma],$$

and that all maximal Cartesian products of K^{sc} are necessarily contained in $\{[\xi,\varsigma] \times [\xi,\varsigma] : (\xi,\varsigma) \in K\}$. An analogous property is no longer true in the vectorial case with $m > 1$, cf. [34, Remark 4.8(c)]. This observation, when applied to the sublevel sets of W, can be viewed as the origin for the structural change that J_W may undergo during relaxation. For illustration, we discuss two explicit examples of L^∞-functionals of multi-well type, one with and one without structure preservation, see Examples 5.5 and 5.6 respectively. The latter relies on what we refer to as the effect of hidden Cartesian squares. It is related to the fact that a union of Cartesian squares

$$\bigcup_{i\in\mathcal{I}} A_i \times A_i$$

with an index set $\mathcal{I}$ and $A_i \times A_i \subset \mathbb{R}^m \times \mathbb{R}^m$ for $i \in \mathcal{I}$ can contain $B \times B$ with $B \neq A_i$ for all $i \in \mathcal{I}$, for details see Definition 2.5. A condition ruling out the existence of such hidden Cartesian squares is the concept of a *basic Cartesian convexification*, as introduced in Definition 3.8. Essentially, it says that every maximal Cartesian square of the Cartesian convex hull of a set in $\mathbb{R}^m \times \mathbb{R}^m$ coincides with the (classical) convexification of a maximal Cartesian square of the original set. In fact, it turns out that this property for the sublevel sets of W gives a full characterization of the cases when relaxation of J_W is representable as in (1.1).

Theorem 1.2 (*Relaxation of Nonlocal Supremal Functionals*). *Let J_W and J_W^{rlx} be as in (1.1) and (1.3) with $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ lower semicontinuous and coercive. If every sublevel set of $\widehat{W}$ has a basic Cartesian convexification according to Definition 3.8, then*

$$J_W^{\mathrm{rlx}} = J_{\widehat{W}^{\times\mathrm{lc}}},$$

where $\widehat{W}^{\times\mathrm{lc}}$ stands for the Cartesian level convex envelope of $\widehat{W}$. Otherwise, there exists no lower semicontinuous and coercive $G : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ such that $J_W^{\mathrm{rlx}} = J_G$ with J_G as in (1.1).

We remark finally that every subset of $\mathbb{R} \times \mathbb{R}$ has a basic Cartesian convexification, so that, in correspondence with (1.4), $J_W^{\mathrm{rlx}} = J_{\widehat{W}^{\times\mathrm{lc}}} = J_{\widehat{W}^{\mathrm{slc}}}$ holds in the one-dimensional setting without further restriction.

To place the results above into a broader context, we establish a relation between nonlocal supremal functionals of the form (1.1), which are defined on essentially bounded functions, with homogeneous double integrals on L^p-spaces, that is,

$$L^p(\Omega; \mathbb{R}^m) \ni u \mapsto \int_\Omega \int_\Omega V(u(x), u(y)) \, dx \, dy \tag{1.5}$$

with $V : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ lower semicontinuous with standard p-growth. The existence theory for minimizers of this class of functionals was developed over the last two decades. In a series of papers, starting with [41] and continued in [13,14,39,42,43], separate convexity of the double integrand has been identified as the necessary and sufficient condition for their weak lower semicontinuity; for a discussion of conditions in the inhomogeneous case, we refer to [13]. When it comes to the relaxation of (1.5), a comprehensive understanding of the problem is still to be developed. Recent counterexamples in [34,37] indicate that even in the one-dimensional case, weak lower semicontinuous envelopes of functionals like (1.5) may belong to a different class in general. Beyond these specific examples, though, neither explicit formulas nor a general characterization of structure preservation under relaxation are currently available. To what extent, the results of this work can shed light on these open questions remains to be explored in future research.

It is well-known that the standard L^p-norm converges to the L^∞-norm as p goes to infinity. More generally, L^p-approximations hold in the local setting, connecting classical integral and supremal functionals approximatively in the limit of diverging power-law exponents, see e.g. [3,19,25,46,47]. They have been used as technical tools, e.g. for proving the existence of absolute minimizers of supremal functionals [8,19], a homogenization result of L^∞-functionals [16] and for the mathematical analysis of dielectric breakdown in composite materials [28].

Stimulated by the local power-law approximations, we provide here a parallel statement in the nonlocal setting. Precisely, we perform a limit process of diverging integrability constants in terms of Γ-convergence (cf. Definition 2.2), which guarantees the convergence of almost-minimizers and infima of double integral functionals to minimizers and minima of a Γ-limit of nonlocal supremal type.

Theorem 1.3 (*Nonlocal L^p-Approximation*). *Let $q > 1$ and $W : \mathbb{R}^m \times \mathbb{R}^m \to [0, \infty)$ be lower semicontinuous with standard q-growth, i.e., there exist constants $c_1, c_2, C > 0$ such that*

$$c_1 |(\xi, \zeta)|^q - c_2 \leq W(\xi, \zeta) \leq C \left(|(\xi, \zeta)|^q + 1 \right) \qquad for \ (\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m. \tag{1.6}$$

For $p \geq q$, let $I_W^p : L^q(\Omega; \mathbb{R}^m) \to [0, \infty]$ be defined as

$$I_W^p(u) = \begin{cases} \left(\displaystyle\int_\Omega \int_\Omega W^p(u(x), u(y)) \, dx \, dy \right)^{\frac{1}{p}} & if \ u \in L^p(\Omega; \mathbb{R}^m), \\ \infty & otherwise. \end{cases}$$

Then, $(I_W^p)_p$ Γ-converges with respect to the weak topology in $L^q(\Omega; \mathbb{R}^m)$ as $p \to \infty$ to the functional $I_W^\infty : L^q(\Omega; \mathbb{R}^m) \to [0, \infty]$ given by

$$I_W^\infty(u) = \begin{cases} J_W^{\text{rlx}}(u) & if \ u \in L^\infty(\Omega; \mathbb{R}^m), \\ \infty & otherwise. \end{cases}$$

Moreover, any sequence $(u_p)_p \subset L^q(\Omega; \mathbb{R}^m)$ with $\sup_{p \geq q} I_W^p(u_p) < \infty$ is relatively sequentially compact in the weak topology of $L^q(\Omega; \mathbb{R}^m)$ and the limits of weakly converging subsequences lie in $L^\infty(\Omega; \mathbb{R}^m)$.

Notice that the previous result does not require any type of convexity assumptions on the double-integrand W, and is new even in the one-dimensional setting. The proof strategy for the compactness and the

Table 1
Comparing properties of double-integrals and nonlocal supremals. Results marked with [∗] are proven in this paper.

	Double integrals	Nonlocal supremals
Invariance under symmetrization	Yes	Yes
Invariance under diagonalization	No	Yes
Necessary and sufficient condition for weak$^{(*)}$ lower semicontinuity	Separate convexity [13,14,39,42]	Cartesian level convexity for $m \geq 1$ [∗] Separate level convexity for $m = 1$ [34]
Structure preservation under relaxation	No [35,37]	Yes for $m = 1$ [34] No for $m > 1$ [∗]
Criteria for structure preservation	Currently still open	Basic Cartesian level convexification [∗]

construction of recovery sequences (cf. Definition 2.2) is rather standard. To obtain the lower bound of the Γ-limit, we take inspiration from [19], while exploiting recent results about the Young measure representation of the weak* closures of solution sets to nonlocal inclusions (see Remark 4.4, and also [35]).

Even though, the two types of functionals – double-integrals and nonlocal supremals – are linked asymptotically in the limit of diverging integrability constants, they show qualitatively very different features, as the table in Table 1 summarizes at a glance.

This paper is organized as follows. After the introduction, we start in Section 2 by fixing notations and gathering some technical tools from asymptotic analysis as well as useful observations about maximal Cartesian squares; other auxiliary results that require more specific terminology have been moved to the Appendix. Section 3 is devoted to the definition and analysis of the new notion of Cartesian (level) convexity, for both sets and functions, as well as the corresponding Cartesian convex hulls and envelopes. In particular, we prove different alternative characterizations, including a Jensen-type formula in the case of functions, discuss some relevant properties, and provide a comparison with standard notions of convexity, precisely with separate convexity. In Section 4, Cartesian convexity is identified as the necessary and sufficient condition for the weak* closedness of the solution set for nonlocal inclusions, and we investigate when the weak* closures give rise to a nonlocal inclusion of the same type, cf. Propositions 4.1 and 4.2, respectively. These findings are then transferred to the context of nonlocal supremal functionals in Section 5, where the proofs of our main results Theorems 1.1 and 1.2 are presented. We illustrate the two scenarios of structure preservation during relaxation and the loss there of with examples of multi-well supremals. Finally, Section 6 contains the proof of Theorem 1.3 and builds a rigorous bridge between the variational theories for nonlocal supremals and double integrals.

2. Preliminaries

We use this section to introduce notation, to recall some well-known concepts, and to prove a few auxiliary results that will be useful in the rest of the article.

2.1. Notation

Let $m \in \mathbb{N}$, $\xi, \zeta \in \mathbb{R}^m$, and $|\cdot|$ be the Euclidean norm in $\mathbb{R}^m$. We denote by $[\xi, \zeta] = \{t\xi + (1-t)\zeta : t \in [0,1]\}$ the segment in $\mathbb{R}^m$ with extreme points ξ, ζ. The set $B_c(a) \subset \mathbb{R}^m$ is the Euclidean ball with center point a of radius c. For a set $A \subset \mathbb{R}^m$, $\mathbb{1}_A$ is the characteristic function of A, meaning $\mathbb{1}_A = 1$ in A and $\mathbb{1}_A = 0$ in $\mathbb{R}^m \setminus A$. By A^{co}, we mean the convex hull of A. The distance between a point $\xi \in \mathbb{R}^m$ and a set $A \subset \mathbb{R}^m$ is $\mathrm{dist}(\xi, A) = \inf_{\zeta \in A} |\xi - \zeta|$. In $\mathbb{R}^m \times \mathbb{R}^m$, we use the norm $|(\xi, \zeta)| = \max\{|\xi|, |\zeta|\}$. Then, for $A, B \subset \mathbb{R}^m$,

$$\mathrm{dist}\big((\xi, \zeta), A \times B\big) = \max\{\mathrm{dist}(\xi, A), \mathrm{dist}(\zeta, B)\}, \tag{2.1}$$

and $d_H(A, B) = \max\{\sup_{\xi \in B} \text{dist}(\xi, A), \sup_{\zeta \in A} \text{dist}(\zeta, B)\}$ is the Hausdorff distance between A and B.

Let $\mathcal{M}(\mathbb{R}^m)$ be the space of positive finite Radon measures and by $\mathcal{P}r(\mathbb{R}^m)$, the subset of probability measures. The support of a measure $\nu \in \mathcal{M}(\mathbb{R}^m)$, in formulas $\text{supp}\,\nu$, consists of all the points $\xi \in \mathbb{R}^m$ such that $\nu(O) > 0$ for every open neighborhood $O \subset \mathbb{R}^m$ of ξ. For $\nu \in \mathcal{M}(\mathbb{R}^m)$, we write

$$[\nu] = \langle \nu, \text{id} \rangle = \int_{\mathbb{R}^m} \xi \, d\nu(\xi)$$

for its barycenter, and recall that the convex hull of any compact set $C \subset \mathbb{R}^m$ can be represented as

$$C^{\text{co}} = \{[\nu] : \nu \in \mathcal{P}r(\mathbb{R}^m), \ \text{supp}\,\nu \subset C\}, \tag{2.2}$$

cf. e.g. [24]. The product measure of $\nu, \mu \in \mathcal{M}(\mathbb{R}^m)$ is denoted by $\nu \otimes \mu$.

Let $1 \leq p \leq \infty$, $\nu \in \mathcal{M}(\mathbb{R}^m)$, and $U \subset \mathbb{R}^n$ be a bounded open set, then $L^p_\nu(U; \mathbb{R}^m)$ stands for the standard Lebesgue spaces. If $1 \leq p < \infty$ and $(u_j)_j \in L^p_\nu(U; \mathbb{R}^m)$ is a weakly convergent sequence with limit $u \in L^p_\nu(U; \mathbb{R}^m)$, we write $u_j \rightharpoonup u$ in $L^p_\nu(U; \mathbb{R}^m)$. Analogously, $u_j \rightharpoonup^* u$ in $L^\infty_\nu(U; \mathbb{R}^m)$ stands for the weak* convergence in $L^\infty_\nu(U; \mathbb{R}^m)$. When $m = 1$, we simply write $L^p_\nu(U)$ and $L^\infty_\nu(U)$, if ν is the Lebesgue measure, $L^p(U; \mathbb{R}^m)$ and $L^\infty(U; \mathbb{R}^m)$.

For $A \subset \mathbb{R}^m$, and $\nu \in \mathcal{M}(\mathbb{R}^m)$, the ν-essential supremum in A of a Borel function $f : \mathbb{R}^m \to \mathbb{R}$ is

$$\nu\text{-}\operatorname*{ess\,sup}_{\xi \in A} f(\xi) = \inf\{c \in \mathbb{R} : f \leq c \ \nu\text{-a.e. in } A\} = \inf_{N \subset A, \nu(N)=0} \sup_{\xi \in A \setminus N} f(\xi);$$

in short-hand, $\nu\text{-}\operatorname{ess\,sup}_A f$, or simply $\nu\text{-}\operatorname{ess\,sup} f$ in case $A = \mathbb{R}^m$. If f is lower semicontinuous, then

$$\nu\text{-}\operatorname*{ess\,sup}_{\xi \in \mathbb{R}^m} f(\xi) = \sup_{\xi \in \text{supp}\,\nu} f(\xi). \tag{2.3}$$

Let X be a normed space and $V : X \to \mathbb{R}$. Then the sublevel set of V at level $c \in \mathbb{R}$ is denoted by

$$L_c(V) = \{x \in X : V(x) \leq c\}.$$

As a simple consequence of this definition, V can be expressed as

$$V(x) = \inf\{c \in \mathbb{R} : x \in L_c(V)\} \quad \text{for } x \in X. \tag{2.4}$$

On the other hand, if a family of monotone increasing sets $(F_c)_{c \in \mathbb{R}} \subset X$ satisfies $\bigcap_{j \in \mathbb{N}} F_{c_j} = F_c$ for any decreasing real sequence $(c_j)_j$ with $c_j \to c$ as $j \to \infty$ and $V(x) = \inf\{c \in \mathbb{R} : x \in F_c\}$ for $x \in X$, then

$$F_c = L_c(V) \quad \text{for every } c \in \mathbb{R}. \tag{2.5}$$

2.2. Tools from asymptotic analysis

We start by recalling some elements from Young measure theory; for more on the topic, see e.g. [26,40]. A map $\nu : U \subset \mathbb{R}^n \to \mathcal{M}(\mathbb{R}^m)$ is called a Young measure if it is essentially bounded and weakly* measurable with $\nu_x = \nu(x) \in \mathcal{P}r(\mathbb{R}^m)$ a.e. $x \in U$; we write $\nu \in L^\infty_w(U; \mathcal{P}r(\mathbb{R}^m))$.

If $(v_j)_j \subset L^p(U; \mathbb{R}^m)$ with $p > 1$ and $\nu \in L^\infty_w(U; \mathcal{P}r(\mathbb{R}^m))$, the sequence $(v_j)_j$ generates the Young measure ν, in formulas,

$$v_j \xrightarrow{\text{YM}} \nu \text{ as } j \to \infty$$

if

$$\lim_{j \to \infty} \int_U f(x) \varphi(v_j(x)) \, dx = \int_U f(x) \int_{\mathbb{R}^m} \varphi(\xi) \, d\nu_x(\xi) \, dx$$

for every $f \in L^1(U)$ and every $\varphi \in C_0(\mathbb{R}^m)$ in the closure of real-valued functions on $\mathbb{R}^m$ with compact support.

The next proposition is a version of the fundamental theorem of Young measures (see [26,40]), as we use it in the proof of Theorem 1.3 in Section 6.

Proposition 2.1. *Let $q > 1$ and $(v_j)_j \subset L^q(U; \mathbb{R}^m)$.*

(i) *If $(v_j)_j$ is bounded in $L^q(U; \mathbb{R}^m)$, then there exists a subsequence (not relabeled) and a Young measure $\nu \in L_w^\infty(U; \mathcal{P}r(\mathbb{R}^m))$ such that $v_j \xrightarrow{YM} \nu$ as $j \to \infty$.*

(ii) *If $(v_j)_j \subset L^q(U; \mathbb{R}^m)$ generates a Young measure ν, then*

$$\liminf_{j \to \infty} \int_U G(x, v_j(x)) \, dx \geq \int_U \int_{\mathbb{R}^m} G(x, \xi) d\nu_x(\xi) \, dx$$

for any nonnegative, normal integrand $G: U \times \mathbb{R}^m \to \mathbb{R}$ such that $(G(\cdot, v_j(\cdot)))_j$ is equi-integrable.

The natural notion for capturing the asymptotic behavior of parameter-dependent variational problems is Γ-convergence of the associated functionals, which we state here in a setting relevant for this paper; a comprehensive introduction to the theory can be found in [21].

Definition 2.2 (Γ-*Convergence*). Let X be a Banach space with separable dual and $I_j : X \to [-\infty, \infty]$ for $j \in \mathbb{N}$ equi-coercive functionals, that is, $I_j \geq \Psi$ for every $j \in \mathbb{N}$ with $\Psi(x) \to \infty$ as $\|x\|_X \to \infty$, where $\| \cdot \|_X$ stands for the norm in X.

Then $(I_j)_j$ Γ-converges to $I : X \to [-\infty, \infty]$ with respect to the weak topology in X, in formulas $I_j \xrightarrow{\Gamma} I$ as $j \to \infty$ or $\Gamma - \lim_{j \to \infty} I_j = I$, if these two conditions hold:

(i) for all $u \in X$ and for every sequence $(u_j)_j \subset X$ with $u_j \rightharpoonup u$ in X,

$$I(u) \leq \liminf_{j \to \infty} I_j(u_j);$$

(ii) for all $u \in X$ with $I(u) < \infty$ there exists a sequence $(u_j)_j \subset X$ such that $u_j \rightharpoonup u$ in X and

$$I(u) \geq \limsup_{j \to \infty} I_j(u_j).$$

The main feature of the Γ-convergence is that it implies the convergence of minimizers of $(I_j)_j$ to minimizers of I, precisely, if u_j is a minimizer for I_j for each $j \in \mathbb{N}$ and u is a cluster point of $(u_j)_j$, then u is a minimizer of I and

$$I(u) = \lim_{j \to \infty} I_j(u_j),$$

cf. [21, Cororally 7.20]. When a family $(I_p)_p$ of functionals $I_p : X \to [-\infty, \infty]$ with a real parameter p is considered, the Γ-convergence $I_p \to I$ as $p \to \infty$ means that $I_{p_j} \to I$ as $j \to \infty$ for every sequence $(p_j)_j \subset \mathbb{R}$ converging to ∞.

We close this section with a well-known classic, the standard L^p-approximation of the L^∞-norm. Let $\nu \in \mathcal{P}r(\mathbb{R}^m)$ and $f \in L_\nu^p(\mathbb{R}^m)$ for $p > 1$. Then, the map $p \mapsto \|f\|_{L_\nu^p(\mathbb{R}^m)}$ is monotone increasing and

$$\lim_{p \to \infty} \|f\|_{L_\nu^p(\mathbb{R}^m)} = \|f\|_{L_\nu^\infty(\mathbb{R}^m)}; \tag{2.6}$$

in particular, $f \in L_\nu^\infty(\mathbb{R}^m)$ if and only if $\sup_{p > 1} \|f\|_{L_\nu^p(\mathbb{R}^m)} < \infty$. For any bounded measurable set $U \subset \mathbb{R}^m$, one obtain as a special case that

$$\lim_{p \to \infty} \|f\|_{L^p(U)} = \|f\|_{L^\infty(U)}. \tag{2.7}$$

2.3. Diagonalization and maximal cartesian squares

A set $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is called symmetric if $(\xi, \zeta) \in E$ if and only if $(\zeta, \xi) \in E$, and following [34], we say that E is diagonal if $(\xi, \xi), (\zeta, \zeta) \in E$ for every $(\xi, \zeta) \in E$. Let $\widehat{E}$ denote the diagonalization and symmetrization of E, that is,

$$\widehat{E} = \{(\xi, \zeta) \in E : \{\xi, \zeta\} \times \{\xi, \zeta\} \subset E\}.$$

These concepts can be carried over to functions by applying the above definitions to their sublevel sets. Precisely, we introduce the symmetrization and diagonalization of $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ as

$$\widehat{W}(\xi, \zeta) = \max \{W(\xi, \zeta), W(\zeta, \xi), W(\xi, \xi), W(\zeta, \zeta)\} \quad \text{for } (\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m. \tag{2.8}$$

It holds then for any $c \in \mathbb{R}$ that

$$L_c(\widehat{W}) = \widehat{L_c(W)}.$$

A closely related notion to diagonality, which is central for our analysis, is that of maximal Cartesian squares. As indicated in the introduction, the identification of certain maximal Cartesian squares allows us to obtain explicit representations for weak* closures on nonlocal inclusions and relaxation formulas for nonlocal supremals. For an arbitrary $E \subset \mathbb{R}^m \times \mathbb{R}^m$, a set $B \times B \subset E$ with $B \subset \mathbb{R}^m$ is said to be a maximal Cartesian square of E if every $A \subset B$ with $A \times A \subset E$ satisfies $A = B$. We refer to the set of all maximal Cartesian squares of E as $\mathcal{P}_E$. Notice that the existence of a maximal Cartesian square in $\mathcal{P}_E$ containing an arbitrary square $A \times A \subset E$ is guaranteed by Zorn's Lemma.

It was shown in [34, Lemma 4.3] that any symmetric and diagonal set $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is equal to the union of its (maximal) Cartesian squares, i.e.,

$$E = \bigcup_{A \times A \subset E} A \times A = \bigcup_{B \times B \in \mathcal{P}_E} B \times B. \tag{2.9}$$

On the other hand, if a subset of $\mathbb{R}^m \times \mathbb{R}^m$ can be written as the union of Cartesian squares, say

$$E = \bigcup_{i \in \mathcal{I}} A_i \times A_i \tag{2.10}$$

with a (possibly uncountable) index set $\mathcal{I}$ and nonempty sets $A_i \subset \mathbb{R}^m$ for $i \in \mathcal{I}$, it is clearly symmetric and diagonal. However, in contrast to what may seem intuitive at first glance, the maximal Cartesian squares of E do not necessarily coincide with Cartesian squares in (2.10), meaning that in general,

$$\mathcal{P}_E \not\subset \bigcup_{i \in \mathcal{I}} \{A_i \times A_i\}.$$

The next example illustrates this effect.

Example 2.3. Consider three sets $A_1, A_2, A_3 \subset \mathbb{R}^m$ with no common element, but pairwise nonempty intersections, that is, $\bigcap_{i=1}^3 A_i = \emptyset$ and $A_i \cap A_j \neq \emptyset$ for $i, j \in \{1, 2, 3\}$ with $i \neq j$; see Fig. 1 for an illustration in two dimensions. With $E = \bigcup_{i=1}^3 A_i \times A_i$ and the nonempty set

$$M := (A_1 \cap A_2) \cup (A_2 \cap A_3) \cup (A_3 \cap A_1), \tag{2.11}$$

it is straightforward to check that $M \times M \subset E$ and $M \neq A_i$ for $i \in \{1, 2, 3\}$. Hence, $\mathcal{P}_E \not\subset \bigcup_{i=1}^3 \{A_i \times A_i\}$. $\triangle$

Before following-up on this example, let us slide in three brief comments.

Remark 2.4. (a) The primary reason for us to study sets of the form (2.10) is their relevance for our discussion of nonlocal supremal functionals of multi-well type in Section 5.2. In fact, the sublevel sets of the associated supremands have exactly this type of structure with a finite index set $\mathcal{I}$.

(b) We point out that the representation of $E \subset \mathbb{R}^m \times \mathbb{R}^m$ as in (2.10) in terms of a union of Cartesian squares is in general not unique, even if one supposes that each A_i with $i \in \mathcal{I}$ is not contained in any A_j with $j \in \mathcal{I} \setminus \{i\}$. As a simple example, take $m = 1$ and

$$E = \bigcup_{(\alpha, \beta) \in [0,1] \times [0,1]} \{\alpha, \beta\} \times \{\alpha, \beta\} = [0, 1] \times [0, 1].$$
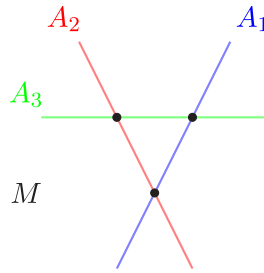
Fig. 1. Illustration of sets A_1, A_2, A_3, and M as in Example 2.3 for $m = 2$.

(c) Let $E \subset \mathbb{R}^m \times \mathbb{R}^m$ be as in (2.10). If $A_i \setminus \bigcup_{j \in \mathcal{I}, j \neq i} A_j \neq \emptyset$ for some $i \in \mathcal{I}$, then $A_i \times A_i \in \mathcal{P}_E$. To show that the Cartesian square $A_i \times A_i \subset E$ is actually maximal, suppose to the contrary that $C_i = A_i \cup \{\gamma\}$ with $\gamma \in A_k \setminus A_i$ for some $k \in \mathcal{I}$ satisfies $C_i \times C_i \subset E$. This implies that (ξ, γ) with $\xi \in A_i \setminus \bigcup_{j \in \mathcal{I}, j \neq i} A_j$ lies neither in $A_i \times A_i$ nor in any other $A_j \times A_j$ with $j \in \mathcal{I} \setminus \{i\}$. Thus, $(\xi, \gamma) \in (C_i \times C_i) \setminus E = \emptyset$, which is a contradiction. $\triangle$

With motivation from Example 2.3, we now introduce the following terminology.

Definition 2.5. Let E be given as in (2.10). Then, $M \times M \subset E$ is called a hidden Cartesian square of E if $M \neq A_i$ for every $i \in \mathcal{I}$. If $M \times M \in \mathcal{P}_E$, we speak of a hidden maximal Cartesian square.

Even though the concept of hidden Cartesian squares is easily accessible in theory, finding ways to identify all of them explicitly for a given set is less obvious. The next result provides two basic conditions, a necessary one and a sufficient one, useful for detecting hidden Cartesian squares.

Lemma 2.6. Let $\mathcal{I}$ be an index set and $E = \bigcup_{i \in \mathcal{I}} A_i \times A_i$ with nonempty $A_i \subset \mathbb{R}^m$ for $i \in \mathcal{I}$. Further, let $M_{\mathcal{I}'} = \bigcup_{i,j \in \mathcal{I}', i \neq j} A_i \cap A_j$ for any $\mathcal{I}' \subset \mathcal{I}$.
  (i) If $A \times A \subset E$, then $A \subset A_i$ for some $i \in \mathcal{I}$ or
$$A \subset M_{\mathcal{I}} = \bigcup_{i,j \in \mathcal{I}, i \neq j} A_i \cap A_j.$$
The latter means that each element of A lies in at least two different sets A_i and A_j with $i, j \in \mathcal{I}$.
  (ii) If $\mathcal{I}' \subset \mathcal{I}$ is such that one can find for any $l, i, j, k \in \mathcal{I}'$ with $i \neq j$ and $k \neq l$ an index $s = s(i, j, k, l) \in \mathcal{I}$ with
$$(A_i \cap A_j) \cup (A_k \cap A_l) \subset A_s, \tag{2.12}$$
then $M_{\mathcal{I}'} \times M_{\mathcal{I}'} \subset E$.

Proof. (i) Arguing by contradiction, let $A \times A \subset E$ be nonempty and assume that for each $i \in \mathcal{I}$ there is $\zeta_i \in A \setminus A_i$, and that there exist $k \in \mathcal{I}$ and $\xi_k \in A$ with the property that $\xi_k \in A_k \setminus \bigcup_{j \in \mathcal{I}, j \neq k} A_j$. Then, $(\xi_k, \zeta_k) \notin A_k \times A_k$ as well as $(\xi_k, \zeta_k) \notin A_j \times A_j$ for all $j \in \mathcal{I} \setminus \{k\}$, and consequently, $(\xi_k, \zeta_k) \notin E$. This contradicts $(\xi_k, \zeta_k) \in A \times A \subset E$.
  (ii) Let $(\xi, \zeta) \in M_{\mathcal{I}'} \times M_{\mathcal{I}'}$. Then, there are $i, j \in \mathcal{I}'$ and $k, l \in \mathcal{I}'$ with $i \neq j$ and $k \neq l$, such that $\xi \in A_i \cap A_j$ and $\zeta \in A_k \cap A_l$. By (2.12), we find $s \in \mathcal{I}$ such that $(\xi, \zeta) \in A_s \times A_s \subset E$. $\square$

Finally, we collect some useful consequences of the previous lemma, in particular, two simple cases where the existence of hidden Cartesian squares can be ruled out and a result for unions of three Cartesian squares. Note that for unions of four (or more) Cartesian squares, a variety of scenarios can occur, as illustrated in Fig. 2.
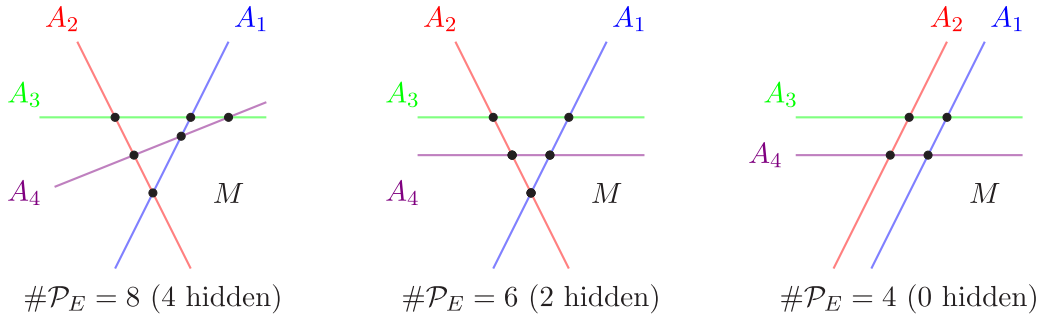
$$\#\mathcal{P}_E = 8 \ (4 \text{ hidden}) \qquad \#\mathcal{P}_E = 6 \ (2 \text{ hidden}) \qquad \#\mathcal{P}_E = 4 \ (0 \text{ hidden})$$

Fig. 2. Number of maximal Cartesian squares for three examples of $E = \bigcup_{i=1}^{4} A_i \times A_i$ with $A_1, A_2, A_3, A_4 \subset \mathbb{R}^2$ and the corresponding $M = M_{\{1,2,3,4\}}$.

Conclusion 2.7. *Let $E = \bigcup_{i \in \mathcal{I}} A_i \times A_i$ with an index set $\mathcal{I}$ and nonempty sets $A_i \subset \mathbb{R}^m$ for $i \in \mathcal{I}$. Suppose that the sets A_i with $i \in \mathcal{I}$ are as large as possible, in the sense that*

$$\bigcup_{i \in \mathcal{I}'} A_i \times \bigcup_{i \in \mathcal{I}'} A_i \not\subset E \tag{2.13}$$

for any index set $\mathcal{I}' \subset \mathcal{I}$ with cardinality greater than one. In particular, (2.13) implies that $A_i \not\subset A_j$ for any $i, j \in \mathcal{I}$ with $i \neq j$. The following statements are immediate consequences of Lemma 2.6.

(a) In case the sets A_i with $i \in \mathcal{I}$ are pairwise disjoint, one has $M_{\mathcal{I}} = \emptyset$, and thus, $\mathcal{P}_E = \bigcup_{i \in \mathcal{I}} \{A_i \times A_i\}$ in view of Lemma 2.6(i) and Remark 2.4(c).

(b) If $\mathcal{I} = \{1, 2\}$, then $M_{\mathcal{I}} = A_1 \cap A_2$ and we obtain $\mathcal{P}_E = \{A_1 \times A_1, A_2 \times A_2\}$, i.e., E does not have any hidden maximal Cartesian squares.

(c) The assumption (2.12) is automatically fulfilled whenever $\#\mathcal{I}' \leq 3$. Therefore, if $\mathcal{I} = \{1, 2, 3\}$, the combination of Lemma 2.6(ii) and (i) implies

$$\mathcal{P}_E \subset \bigcup_{i=1}^{3} \{A_i \times A_i\} \cup \{M_{\mathcal{I}} \times M_{\mathcal{I}}\}.$$

This shows that E can have at most one hidden maximal Cartesian square. For an example with $\#\mathcal{P}_E = 4$, we refer back to Example 2.3 and Fig. 1.

3. Cartesian convexity

This section revolves around our new notion of convexity for sets and functions, called Cartesian (level) convexity. As pointed out in the introduction, it plays a crucial role in characterizing the lower semicontinuity of nonlocal supremal and their relaxations.

3.1. Cartesian convexity of sets

Recall that a set $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is separately convex (with vector components) if

$$[\xi_1, \xi_2] \times [\zeta_1, \zeta_2] \subset E \qquad \text{for all } (\xi_1, \zeta_1), (\xi_2, \zeta_2) \in E \text{ such that } \xi_1 = \xi_2 \text{ or } \zeta_1 = \zeta_2.$$

The separately convex hull of E is denoted by E^{sc}.

With the motivation to introduce a suitable (in general) weaker concept of convexity for subsets of $\mathbb{R}^m \times \mathbb{R}^m$, we give the following definition.

Definition 3.1 (*Cartesian Convexity of Sets*). We say that $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is Cartesian convex if

$$[\xi_1, \xi_2] \times [\zeta_1, \zeta_2] \subset E$$

for all $(\xi_1, \zeta_1), (\xi_2, \zeta_2) \in E$ such that $\{\xi_1, \xi_2, \zeta_1, \zeta_2\} \times \{\xi_1, \xi_2, \zeta_1, \zeta_2\} \subset E$ and $\xi_1 = \xi_2$ or $\zeta_1 = \zeta_2$.

The next lemma states useful alternative characterizations of Cartesian convexity.

Lemma 3.2. *For any $E \subset \mathbb{R}^m \times \mathbb{R}^m$, the following conditions are equivalent:*

(i) E is Cartesian convex;
(ii) $(A \times A)^{\mathrm{co}} = A^{\mathrm{co}} \times A^{\mathrm{co}} \subset E$ for all $A \subset \mathbb{R}^m$ with $A \times A \subset E$;
(iii) any maximal Cartesian square of E is convex.

If E is symmetric and diagonal, then $(i) - (iii)$ are also equivalent to

(iv) $E = \bigcup_{A \times A \subset E} A^{\mathrm{co}} \times A^{\mathrm{co}} = \bigcup_{B \times B \in \mathcal{P}_E} B^{\mathrm{co}} \times B^{\mathrm{co}}$.

Proof. As for $(i) \Rightarrow (ii)$, let $A \times A \subset E$ and consider a maximal Cartesian square $B \times B \in \mathcal{P}_E$ with $A \subset B$. We claim that B is convex. To show this, take $\xi_1, \xi_2 \in B$ and observe that (i) implies $([\xi_1, \xi_2] \times \{\alpha\}) \cup (\{\beta\} \times [\xi_1, \xi_2]) \subset E$ for any $\alpha, \beta \in B$, or in other words,

$$([\xi_1, \xi_2] \times B) \cup (B \times [\xi_1, \xi_2]) \subset E. \tag{3.1}$$

On the other hand, one can deduce from the Cartesian convexity of E that

$$[\xi_1, \xi_2] \times [\xi_1, \xi_2] \subset E;$$

indeed, for $\alpha \in [\xi_1, \xi_2]$ it follows from (3.1) and (i) that $([\xi_1, \xi_2] \times \{\alpha\}) \cup (\{\alpha\} \times [\xi_1, \xi_2]) \subset E$.

Summing up, we have that

$$([\xi_1, \xi_2] \cup B) \times ([\xi_1, \xi_2] \cup B) \subset E.$$

Due to the maximality property of $B \times B$, this yields $[\xi_1, \xi_2] \subset B$, proving the convexity of B. Consequently, $A^{\mathrm{co}} \subset B$ and therefore, $(A \times A)^{\mathrm{co}} = A^{\mathrm{co}} \times A^{\mathrm{co}} \subset B \times B \subset E$, as stated in (ii).

The remaining implications are straightforward to verify. Indeed, $(ii) \Rightarrow (i)$ is a direct consequence of Definition 3.1, the equivalence $(ii) \Leftrightarrow (iii)$ is clear considering the definition of maximal Cartesian squares, and $(ii) \Leftrightarrow (iv)$ holds in light of (2.9). $\square$

Remark 3.3 (*Comparison with Separate Convexity*). Let $E \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric and diagonal. It is evident that separate convexity is sufficient for Cartesian convexity. Whether it is also necessary depends on the dimension m.

(a) In the scalar setting, where $m = 1$, the set E is Cartesian convex if and only if it is separately convex. Indeed, if $E \subset \mathbb{R} \times \mathbb{R}$ is Cartesian convex, we invoke [34, Lemma 4.7] to find that the separately convex hull of E can be expressed as the union of all symmetric squares with corners in the points of E, formally,

$$E^{\mathrm{sc}} = \bigcup_{(\alpha, \beta) \in E} [\alpha, \beta] \times [\alpha, \beta].$$

Then, along with Lemma 3.2(iv),

$$E^{\mathrm{sc}} = \bigcup_{\{\alpha, \beta\} \times \{\alpha, \beta\} \subset E} \{\alpha, \beta\}^{\mathrm{co}} \times \{\alpha, \beta\}^{\mathrm{co}} \subset \bigcup_{A \times A \subset E} A^{\mathrm{co}} \times A^{\mathrm{co}} = E,$$

which shows the separate convexity of E.

(b) If $m > 1$, separate convexity is strictly stronger than Cartesian convexity, as this example illustrates. Consider the symmetric and diagonal set

$$E = (A_1 \times A_1) \cup (A_2 \times A_2),$$

where $A_1, A_2 \subset \mathbb{R}^m$ are two convex sets with nonempty intersection such that $(A_1 \cup A_2)^{\mathrm{co}} \setminus (A_1 \cup A_2) \neq \emptyset$. Since $\mathcal{P}_E = \{A_1 \times A_1, A_2 \times A_2\}$ according to Conclusion 2.7(b), the set E satisfies the condition in Lemma 3.2(iii), and is thus Cartesian convex. To see that E is not separately convex, we argue that

$$E^{\mathrm{sc}} = E \cup \left[(A_1 \cap A_2) \times (A_1 \cup A_2)^{\mathrm{co}}\right] \cup \left[(A_1 \cup A_2)^{\mathrm{co}} \times (A_1 \cap A_2)\right] \neq E,$$

cf. [34, Remark 4.6 (b)]. Observe that E^{sc} and E do coincide after diagonalization, though, i.e., $\widehat{E^{\mathrm{sc}}} = E$. △

3.2. Cartesian convexification

If $E \subset \mathbb{R}^m \times \mathbb{R}^m$ is not Cartesian convex, one may consider the smallest Cartesian convex set containing E, which we call the Cartesian convex hull of E and denote by $E^{\times \mathrm{c}}$; precisely,

$$E^{\times \mathrm{c}} = \bigcap_{F \in \mathcal{F}_E} F \qquad \text{with } \mathcal{F}_E = \{F \subset \mathbb{R}^m \times \mathbb{R}^m : E \subset F, F \text{ is Cartesian convex}\}.$$

Note that $E^{\times \mathrm{c}}$ is well-defined by the simple observation that intersections (even uncountable) of Cartesian convex sets are again Cartesian convex. If E is symmetric and diagonal, so is $E^{\times \mathrm{c}}$.

In the spirit of [20, Theorem 7.17] on the representation of separately convex hulls, we prove the following analogue for Cartesian convex hulls of symmetric and diagonal sets.

Lemma 3.4 (*Representation of $E^{\times \mathrm{c}}$*). *Let $E \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric and diagonal. Then,*

$$E^{\times \mathrm{c}} = \bigcup_{k \in \mathbb{N}} E_k^{\times \mathrm{c}}$$

with inductively defined sets $E_0^{\times \mathrm{c}} := E$ and

$$E_k^{\times \mathrm{c}} := \bigcup_{A \times A \subset E_{k-1}^{\times \mathrm{c}}} A^{\mathrm{co}} \times A^{\mathrm{co}} = \bigcup_{B \times B \in \mathcal{P}_{E_{k-1}^{\times \mathrm{c}}}} B^{\mathrm{co}} \times B^{\mathrm{co}} \qquad \text{for } k \in \mathbb{N}. \tag{3.2}$$

Proof. The argument is quite standard. Applying Lemma 3.2(ii) iteratively shows that $E^{\times \mathrm{c}}$ contains all the nested sets $E_k^{\times \mathrm{c}}$ for $k \in \mathbb{N}$. The reverse inclusion follows from the Cartesian convexity of $\bigcup_{k \in \mathbb{N}} E_k^{\times \mathrm{c}} =: \widetilde{E}$. Indeed, given $(\xi_1, \zeta), (\xi_2, \zeta) \in \widetilde{E}$ with $\{\xi_1, \xi_2, \zeta\} \times \{\xi_1, \xi_2, \zeta\} \subset \widetilde{E}$, one can find $k \in \mathbb{N}$ such that $\{\xi_1, \xi_2, \zeta\} \times \{\xi_1, \xi_2, \zeta\} \subset E_k^{\times \mathrm{c}}$, and hence, $\{\xi_1, \xi_2, \zeta\}^{\mathrm{co}} \times \{\xi_1, \xi_2, \zeta\}^{\mathrm{co}} \subset E_{k+1}^{\times \mathrm{c}} \subset \widetilde{E}$. In particular, $[\xi_1, \xi_2] \times \{\zeta\} \subset \widetilde{E}$, which concludes the proof. □

Next, we determine explicitly the Cartesian convex hulls for two examples. The second one shows in particular that $E_1^{\times \mathrm{c}}$ is not Cartesian convex in general.

Example 3.5. Let $E = \bigcup_{i=1}^3 A_i \times A_i$ with $A_i \subset \mathbb{R}^m$ for $i = 1, 2, 3$.

(a) Suppose that A_i are convex and $A_i \cap A_j \neq \emptyset$ for $i \neq j$. With $M = \bigcup_{i,j=1,2,3,i \neq j} A_i \cap A_j$, we obtain $\mathcal{P}_E \subset \bigcup_{i=1}^3 \{A_i \times A_i\} \cup \{M \times M\}$ by Conclusion 2.7 (c), and thus,

$$E_1^{\times \mathrm{c}} = \bigcup_{i=1}^3 (A_i^{\mathrm{co}} \times A_i^{\mathrm{co}}) \cup (M^{\mathrm{co}} \times M^{\mathrm{co}}). \tag{3.3}$$
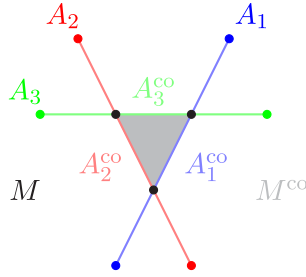
Fig. 3. Illustration of Example 3.5 (b) for $m = 2$.

Lemma 2.6(i) then shows that $E_1^{\times c}$ as represented in (3.3) does not have any hidden squares, so that

$$\mathcal{P}_{E_1^{\times c}} \subset \bigcup_{i=1}^{3} \{A_i^{\mathrm{co}} \times A_i^{\mathrm{co}}\} \cup \{M^{\mathrm{co}} \times M^{\mathrm{co}}\} = \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_E\}.$$

In view of Lemma 3.4, this implies $E^{\times c} = E_1^{\times c}$.

(b) Suppose now that A_i are pairwise disjoint such that the convex hulls A_i^{co} have pairwise nonempty intersection and satisfy $M \setminus A_i^{\mathrm{co}} \neq \emptyset$ for $i = 1, 2, 3$ with $M = \bigcup_{i,j=1,2,3,i\neq j} A_i^{\mathrm{co}} \cap A_j^{\mathrm{co}}$.

Then, $\mathcal{P}_E = \bigcup_{i=1}^{3} \{A_i \times A_i\}$ by Conclusion 2.7 (a) gives

$$E_1^{\times c} = \bigcup_{i=1}^{3} A_i^{\mathrm{co}} \times A_i^{\mathrm{co}},$$

and we infer from Conclusion 2.7 (c) that $E_1^{\times c}$ contains exactly four maximal Cartesian squares, namely, $\mathcal{P}_{E_1^{\times c}} = \bigcup_{i=1}^{3} \{A_i^{\mathrm{co}} \times A_i^{\mathrm{co}}\} \cup \{M \times M\}$. Hence,

$$E_2^{\times c} = E_1^{\times c} \cup (M^{\mathrm{co}} \times M^{\mathrm{co}}).$$

Since $E^{\times c} \supset E_2^{\times c} \supsetneq E_1^{\times c}$, we conclude that $E_1^{\times c}$ cannot be Cartesian convex. In fact, it holds that $E^{\times c} = E_2^{\times c}$, which we can see by applying (a) to the convexifications of the sets A_i, see Fig. 3.   △

It is well-known that the convex hull of a compact set is again compact. However, this is no longer true for the separately convex hull, see [7, Example 2.2] and [32, Example 2.1] (or [20, Remark 8.7]). Whether Cartesian convexification preserves compactness is currently an open question. Nevertheless, the following result holds.

Lemma 3.6. *Let $K \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric, diagonal, and compact. Then, $K_k^{\times c}$ is compact for any $k \in \mathbb{N}$.*

Proof. We prove the statement first for $k = 1$. Since $K_1^{\times c} \subset K^{\mathrm{co}}$, it is clear that $K_1^{\times c}$ is bounded. To see that it is also closed, let $(\xi_j, \zeta_j) \subset K_1^{\times c}$ for $j \in \mathbb{N}$ and suppose that $(\xi_j, \zeta_j) \to (\xi, \zeta)$ as $j \to \infty$ for some $(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m$. We aim to show that $(\xi, \zeta) \in K_1^{\times c}$.

In light of (3.2), there exists a sequence of compact sets $(B_j)_j \subset \mathbb{R}^m$ such that $B_j \times B_j \in \mathcal{P}_K$ and $(\xi_j, \zeta_j) \in B_j^{\mathrm{co}} \times B_j^{\mathrm{co}}$ for all j. From Lemma A.1, we can infer the existence of a compact $B \subset \mathbb{R}^m$ with $B \times B \subset K$ and

$$d_H(B_j^{\mathrm{co}} \times B_j^{\mathrm{co}}, B^{\mathrm{co}} \times B^{\mathrm{co}}) \leq d_H(B_j^{\mathrm{co}}, B^{\mathrm{co}}) \to 0 \quad \text{as } j \to \infty,$$

if necessary, after passing to a (non-relabeled) subsequence. Hence,

$$0 = \lim_{j \to \infty} d_H(B_j^{\mathrm{co}} \times B_j^{\mathrm{co}}, B^{\mathrm{co}} \times B^{\mathrm{co}})$$
$$\geq \lim_{j \to \infty} \mathrm{dist}((\xi_j, \zeta_j), B^{\mathrm{co}} \times B^{\mathrm{co}}) = \mathrm{dist}((\xi, \zeta), B^{\mathrm{co}} \times B^{\mathrm{co}}),$$

which implies $(\xi, \zeta) \in B^{\mathrm{co}} \times B^{\mathrm{co}} \subset K_1^{\times c}$, recalling that $B \times B$ is a subset of K.

The compactness of $K_k^{\times c}$ with $k \in \mathbb{N}$ can be obtained inductively by repeating the above argument. $\quad\square$

Remark 3.7. In (2.2), we state the well-known characterization of convex hulls of compact sets via barycenters of probability measures. An analogous representation formula holds for $K_1^{\times c}$ with $K \subset \mathbb{R}^m \times \mathbb{R}^m$ symmetric, diagonal, and compact. Precisely,

$$K_1^{\times c} = \{[\Lambda] : \Lambda = \nu \otimes \mu \text{ with } \nu, \mu \in \mathcal{P}r(\mathbb{R}^m) \text{ such that } \mathrm{supp}(\nu \otimes \mu) \subset B \times B \text{ for some } B \times B \in \mathcal{P}_K\}.$$

This is straightforward to verify, given (2.2) and the observation that B is compact for every $B \times B \in \mathcal{P}_K$. $\quad\triangle$

Next, we introduce a condition regarding the Cartesian convexification of symmetric and diagonal sets that turns out to be the key to characterizing structure preservation of both nonlocal inclusions under L^∞-weakly* closure and nonlocal supremals under relaxation in the vectorial setting, cf. Propositions 4.2 and 5.2.

Definition 3.8 (*Basic Cartesian Convexification*). Let $E \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric and diagonal. We say that E has a basic Cartesian convexification, or a basic Cartesian convex hull, if

$$\mathcal{P}_{E^{\times c}} \subset \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_E\}. \tag{3.4}$$

Let us make a few basic observations and discuss some useful properties of this notion.

Remark 3.9. Let $E \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric and diagonal.

(a) Any Cartesian convex set E has a basic Cartesian convexification. This is due to

$$\mathcal{P}_{E^{\times c}} = \mathcal{P}_E = \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_E\},$$

where the last identity exploits the maximality property of the elements of $\mathcal{P}_E$ as well as Lemma 3.2.

(b) Suppose that E has a basic Cartesian convexification, then

$$E^{\times c} = E_1^{\times c}. \tag{3.5}$$

Indeed, we deduce from (2.9) and (3.4) that

$$E^{\times c} = \bigcup_{A \times A \in \mathcal{P}_{E^{\times c}}} A \times A \subset \bigcup_{B \times B \in \mathcal{P}_E} B^{\mathrm{co}} \times B^{\mathrm{co}} = E_1^{\times c} \subset E^{\times c}.$$

If E is additionally compact, then $E^{\times c}$ is also compact by Lemma 3.6.

(c) Note that (3.4) can be replaced by

$$\mathcal{P}_{E_1^{\times c}} \subset \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_E\}. \tag{3.6}$$

To see this, recall Lemma 3.4 and apply (3.6) iteratively to obtain $E_1^{\times c} = E_2^{\times c} = \cdots = E_k^{\times c}$ for every $k \in \mathbb{N}$, and thus, $E^{\times c} = E_1^{\times c}$.

(d) In the scalar setting and for compact sets, the condition of having a basic Cartesian convexification is no restriction. If $E \subset \mathbb{R} \times \mathbb{R}$ is compact, we know $E^{\times \mathrm{c}} = E^{\mathrm{sc}}$ by Remark 3.3 (a), and therefore, due to [34, Lemma 4.7],

$$\mathcal{P}_{E^{\times \mathrm{c}}} = \mathcal{P}_{E^{\mathrm{sc}}} \subset \{[\alpha, \beta] \times [\alpha, \beta] : (\alpha, \beta) \in E\} \subset \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \subset E\}.$$

In particular, it follows then from (b) that $E^{\times \mathrm{c}} = E_1^{\times \mathrm{c}}$.

(e) An example of a set that fails to satisfy the necessary condition (3.5) and does therefore not possess a basic Cartesian convexification, has been discussed already in Example 3.5 (b). △

Finally, let us comment on the relation between $E^{\times \mathrm{c}}$ and $\widehat{E^{\mathrm{sc}}}$. Even though we know that $\widehat{E^{\mathrm{sc}}}$ is Cartesian convex by Lemma 3.2 and clearly, $E^{\times \mathrm{c}} \subset \widehat{E^{\mathrm{sc}}}$, it is currently still open whether the identity

$$E^{\times \mathrm{c}} = \widehat{E^{\mathrm{sc}}}$$

holds in general in the vectorial case. For $m = 1$, it follows immediately from Remark 3.3 (a), and it is also satisfied for the higher-dimensional example in Remark 3.3 (b). Besides, the question whether $\widehat{E^{\mathrm{sc}}}$ and $E^{\times \mathrm{c}}$ are always compact provided E is compact remains to be resolved.

3.3. Cartesian level convexity of functions

Our new notion of convexity for sets can be carried over to functions by requiring that their sublevel sets are Cartesian convex.

Definition 3.10 (*Cartesian Level Convexity*). A function $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is called Cartesian level convex if all sublevel sets of W, i.e., $L_c(W)$ with $c \in \mathbb{R}$, are Cartesian convex.

As we prove next, there is a (generalized) Jensen-type inequality for Cartesian level convex functions.

Proposition 3.11. *Let $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be lower semicontinuous. Then, the following statements are equivalent:*

(i) W is Cartesian level convex;
(ii) for any $\xi_1, \xi_2, \zeta_1, \zeta_2 \in \mathbb{R}^m$, it holds that

$$
\begin{aligned}
W(\alpha, \beta) &\leq \max_{i,j \in \{1,2\}} \{W(\xi_i, \zeta_j), W(\zeta_i, \xi_j), W(\xi_i, \xi_j), W(\zeta_i, \zeta_j)\} \\
&= \max_{(\xi, \zeta) \in \{\xi_1, \xi_2, \zeta_1, \zeta_2\} \times \{\xi_1, \xi_2, \zeta_1, \zeta_2\}} W(\xi, \zeta)
\end{aligned}
\tag{3.7}
$$

for all $(\alpha, \beta) \in [\xi_1, \xi_2] \times [\zeta_1, \zeta_2]$;
(iii) for every $\nu, \mu \in \mathcal{P}r(\mathbb{R}^m)$,

$$W([\nu], [\mu]) \leq \max \left\{ (\nu \otimes \mu)\text{-}\mathrm{ess\,sup}\, W, (\mu \otimes \nu)\text{-}\mathrm{ess\,sup}\, W, (\nu \otimes \nu)\text{-}\mathrm{ess\,sup}\, W, (\mu \otimes \mu)\text{-}\mathrm{ess\,sup}\, W \right\}. \tag{3.8}$$

Proof. $(i) \Rightarrow (ii)$ Let $\xi_1, \xi_2, \zeta_1, \zeta_2 \in \mathbb{R}^m$ and $(\alpha, \beta) \in [\xi_1, \xi_2] \times [\zeta_1, \zeta_2]$. Moreover, let c be the right-hand side of (3.7). Setting $A := \{\xi_1, \xi_2, \zeta_1, \zeta_2\}$, it holds that

$$A \times A \subset L_c(W).$$

It follows then by the Cartesian convexity of $L_c(W)$ that

$$[\xi_1, \xi_2] \times [\zeta_1, \zeta_2] \subset A^{\mathrm{co}} \times A^{\mathrm{co}} \subset L_c(W),$$

which implies $W(\alpha, \beta) \leq c$.

$(ii) \Rightarrow (i)$ Let $c \in \mathbb{R}$ and $(\xi_1, \zeta_1), (\xi_2, \zeta_2) \in L_c(W)$ be such that $\{\xi_1, \xi_2, \zeta_1, \zeta_2\} \times \{\xi_1, \xi_2, \zeta_1, \zeta_2\} \subset L_c(W)$. We then infer from (3.7) that $W(\alpha, \beta) \leq c$ for all $(\alpha, \beta) \in [\xi_1, \xi_2] \times [\zeta_1, \zeta_2]$, and hence, $[\xi_1, \xi_2] \times [\zeta_1, \zeta_2] \subset L_c(W)$.

$(iii) \Rightarrow (ii)$ If $\alpha = s\xi_1 + (1-s)\xi_2$ and $\beta = t\zeta_1 + (1-t)\zeta_2$ for $s, t \in [0,1]$ and $\xi_1, \xi_2, \zeta_1, \zeta_2 \in \mathbb{R}^m$, then (ii) follows directly from the choice of the probability measures

$$\nu = s\delta_{\xi_1} + (1-s)\delta_{\xi_2} \quad \text{and} \quad \mu = t\delta_{\zeta_1} + (1-t)\delta_{\zeta_2}.$$

$(ii) \Rightarrow (iii)$ Let $\nu, \mu \in \mathcal{P}r(\mathbb{R}^m)$ and denote the right-hand side of (3.8) by c. Since $[\nu] \in (\mathrm{supp}\,\nu)^{\mathrm{co}}$ and $[\mu] \in (\mathrm{supp}\,\mu)^{\mathrm{co}}$, by Caratheodory's formula for convex hulls [20, Theorem 2.13], there exist $\xi_i \in \mathrm{supp}\,\nu$, $\zeta_j \in \mathrm{supp}\,\mu$ and $t_i, s_j \in [0,1]$ for $i, j = 1, \ldots, m+1$ such that $\sum_{i=1}^{m+1} t_i = 1 = \sum_{j=1}^{m+1} s_j$ and

$$([\nu], [\mu]) = \left(\sum_{i=1}^{m+1} t_i \xi_i \,, \sum_{j=1}^{m+1} s_j \zeta_j \right).$$

We will show next that

$$A \times A \subset L_c(W) \qquad \text{with } A = \{\xi_1, \ldots, \xi_{m+1}, \zeta_1, \ldots, \zeta_{m+1}\}. \tag{3.9}$$

Indeed, for all i, j, one obtains in view of $(\xi_i, \zeta_j) \in \mathrm{supp}\,\nu \times \mathrm{supp}\,\mu$ and the identity (2.3) that

$$W(\xi_i, \zeta_j) \leq \sup_{\mathrm{supp}(\nu \otimes \mu)} W = (\nu \otimes \mu)\text{-}\mathrm{ess\,sup}\,W \leq c,$$

and therefore $(\xi_i, \zeta_j) \in L_c(W)$; an analogous argument shows that $(\zeta_i, \xi_j), (\xi_i, \xi_j), (\zeta_i, \zeta_j) \in L_c(W)$. This implies (3.9).

Since the sublevel set $L_c(W)$ is Cartesian convex, we can invoke Lemma 3.2 to conclude that

$$([\nu], [\mu]) \in A^{\mathrm{co}} \times A^{\mathrm{co}} \subset L_c(W). \quad \square$$

In parallel to Section 3.1, let us comment briefly on the relation between Cartesian and separate level convexity. Recall that a function $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is separately level convex if all sublevel sets of W are separately convex.

Remark 3.12 (*Comparison with Separate Level Convexity*). The equivalence of these three conditions for a lower semicontinuous $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is known from [34, Lemma 3.5]:

(i) W is separately level convex;
(ii) for any $\xi_1, \xi_2, \zeta_1, \zeta_2 \in \mathbb{R}^m$, it holds that

$$W(\alpha, \beta) \leq \max_{i,j \in \{1,2\}} W(\xi_i, \zeta_j) = \max_{(\xi, \zeta) \in \{\xi_1, \xi_2\} \times \{\zeta_1, \zeta_2\}} W(\xi, \zeta) \tag{3.10}$$

for all $(\alpha, \beta) \in [\xi_1, \xi_2] \times [\zeta_1, \zeta_2]$;
(iii) for every $\nu, \mu \in \mathcal{P}r(\mathbb{R}^m)$,

$$W([\nu], [\mu]) \leq (\nu \otimes \mu)\text{-}\mathrm{ess\,sup}\,W. \tag{3.11}$$

It is apparent that (ii) and (iii) imply Proposition 3.11(ii) and (iii), which reflects that any separately level convex function is also Cartesian level convex. When $m = 1$ and W is also symmetric and diagonal, the notions of Cartesian and separate level convexity are identical in light of Remark 3.3 (a). Therefore, the Jensen's formulas (3.7) and (3.10) as well as (3.8) and (3.11) coincide in that case. $\triangle$

When W fails to be Cartesian level convex, the corresponding envelope defined by

$$W^{\times \mathrm{lc}}(\xi, \zeta) := \sup\{V(\xi, \zeta) : V : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \text{ is Cartesian level convex and } V \leq W\}, \tag{3.12}$$

for $(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m$ provides the largest Cartesian level convex function not exceeding W. The next lemma, which is based on the results of Section 3.2, implies under suitable assumptions on W and under consideration of (2.4) a natural representation of $W^{\times \mathrm{lc}}$, namely,

$$W^{\times \mathrm{lc}}(\xi, \zeta) = \inf\{c \in \mathbb{R} : (\xi, \zeta) \in L_c(W)^{\times \mathrm{c}}\} \qquad \text{for } (\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m.$$

Lemma 3.13. *Let $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be symmetric, diagonal, lower semicontinuous and coercive. Then, $W^{\times \mathrm{lc}}$ is also symmetric, diagonal, and coercive, and satisfies for any $c \in \mathbb{R}$ that*

$$L_c(W^{\times \mathrm{lc}}) = L_c(W)^{\times \mathrm{c}}. \tag{3.13}$$

If all sublevel sets of W admit a basic Cartesian convexification in the sense of Definition 3.8, then $W^{\times \mathrm{lc}}$ is also lower semicontinuous.

Proof. To show (3.13), let $c \in \mathbb{R}$ be fixed. Since $W^{\times \mathrm{lc}} \leq W$, and thus, $L_c(W) \subset L_c(W^{\times \mathrm{lc}})$, one has

$$L_c(W)^{\times \mathrm{c}} \subset L_c(W^{\times \mathrm{lc}})^{\times \mathrm{c}} = L_c(W^{\times \mathrm{lc}}),$$

where the last identity follows from the Cartesian level convexity of $W^{\times \mathrm{lc}}$.

For the reverse inclusion, consider the auxiliary function

$$\widetilde{W}(\xi, \zeta) := \inf\{c \in \mathbb{R} : (\xi, \zeta) \in L_c(W)^{\times \mathrm{c}}\} \quad \text{for } (\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m.$$

Since Lemma A.2 gives

$$\bigcap_{j \in \mathbb{N}} L_{c + \frac{1}{j}}(W)^{\times \mathrm{c}} = L_c(W)^{\times \mathrm{c}},$$

it follows with (2.5) that $L_c(\widetilde{W}) = L_c(W)^{\times \mathrm{c}}$. Consequently, $\widetilde{W}$ is Cartesian level convex and therefore, $W^{\times \mathrm{lc}} \geq \widetilde{W}$. This implies

$$L_c(W^{\times \mathrm{lc}}) \subset L_c(\widetilde{W}) = L_c(W)^{\times \mathrm{c}},$$

completing the proof of (3.13).

The coercivity and lower semicontinuity of $W^{\times \mathrm{lc}}$ results from the observation that its level sets are bounded and closed, respectively. The latter holds under the additional assumption of a basic Cartesian convex hull due to Remark 3.9 (b) and Lemma 3.6. □

Finally, we relate $W^{\times \mathrm{lc}}$ with the separately level convex envelope of W, defined in analogy to (3.12) and denoted by W^{slc}.

Remark 3.14 (*Comparison with Separate Level Convexification*). Let $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be symmetric, diagonal, and lower semicontinuous. Then, $W^{\times \mathrm{lc}} \geq \widehat{W^{\mathrm{slc}}}$ is a consequence of Remark 3.12 along with the symmetry and diagonality of $W^{\times \mathrm{lc}}$ (cf. Lemma 3.13). The validity of the identity

$$W^{\times \mathrm{lc}} = \widehat{W^{\mathrm{slc}}}$$

remains an open question, but we can show that $\widehat{W^{\mathrm{slc}}}$ is Cartesian level convex by the generalized Jensen's inequality in Proposition 3.11(*iii*). Indeed, for any $\nu, \mu \in \mathcal{P}r(\mathbb{R}^m)$,

$$
\begin{aligned}
\max\big\{&(\nu \otimes \mu)\text{-}\operatorname{ess\,sup} \widehat{W^{\mathrm{slc}}}, (\mu \otimes \nu)\text{-}\operatorname{ess\,sup} \widehat{W^{\mathrm{slc}}}, (\nu \otimes \nu)\text{-}\operatorname{ess\,sup} \widehat{W^{\mathrm{slc}}}, (\mu \otimes \mu)\text{-}\operatorname{ess\,sup} \widehat{W^{\mathrm{slc}}}\big\} \\
&\geq \max\big\{(\nu \otimes \mu)\text{-}\operatorname{ess\,sup} W^{\mathrm{slc}}, (\mu \otimes \nu)\text{-}\operatorname{ess\,sup} W^{\mathrm{slc}}, (\nu \otimes \nu)\text{-}\operatorname{ess\,sup} W^{\mathrm{slc}}, (\mu \otimes \mu)\text{-}\operatorname{ess\,sup} W^{\mathrm{slc}}\big\} \\
&\geq \max\{W^{\mathrm{slc}}([\nu], [\mu]), W^{\mathrm{slc}}([\mu], [\nu]), W^{\mathrm{slc}}([\nu], [\nu]), W^{\mathrm{slc}}([\mu], [\mu]), \} \\
&= \widehat{W^{\mathrm{slc}}}([\nu], [\mu]),
\end{aligned}
$$

where we have exploited (2.8) along with the observations of Remark 3.12. △

4. Weak* closures of nonlocal inclusions

The relaxation of nonlocal L^∞-functionals is closely interlinked with the characterization of weak* limits of sequences satisfying nonlocal inclusions. Building on previous findings in [34,35], we approach this problem from a different angle and investigate the issue of structure preservation in the general vectorial setting.

For $E \subset \mathbb{R}^m \times \mathbb{R}^m$, we introduce the set of solutions $u \in L^\infty(\Omega; \mathbb{R}^m)$ to the nonlocal inclusion $(u(x), u(y)) \in E$ for a.e. $(x,y) \in \Omega \times \Omega$,

$$\mathcal{A}_E = \{u \in L^\infty(\Omega; \mathbb{R}^m) : (u(x), u(y)) \in E \text{ for a.e. } (x,y) \in \Omega \times \Omega\}.$$

In [34], for a compact set $K \subset \mathbb{R}^m \times \mathbb{R}^m$, it was proved that

$$\mathcal{A}_K = \bigcup_{B \times B \in \mathcal{P}_K} \mathcal{A}_{B \times B} = \bigcup_{B \times B \in \mathcal{P}_K} L^\infty(\Omega; B), \tag{4.1}$$

where $L^\infty(\Omega; B)$ denotes the essentially bounded Lebesgue measurable functions from Ω, taking values in B. Moreover, the inclusion $\mathcal{A}_K$ is invariant under symmetrization and diagonalization of K, i.e.,

$$\mathcal{A}_K = \mathcal{A}_{\widehat{K}},$$

and that the order principle

$$\mathcal{A}_K \subset \mathcal{A}_L \qquad \text{if and only if} \qquad \widehat{K} \subset \widehat{L} \tag{4.2}$$

holds for all $K, L \subset \mathbb{R}^m \times \mathbb{R}^m$ compact.

We define the weak* closure of $\mathcal{A}_E$ as

$$\mathcal{A}_E^\infty := \{u \in L^\infty(\Omega; \mathbb{R}^m) : u_j \rightharpoonup^* u \text{ in } L^\infty(\Omega; \mathbb{R}^m) \text{ with } (u_j)_j \subset \mathcal{A}_E\}$$

in analogy to [34]. If $K \subset \mathbb{R}^m \times \mathbb{R}^m$ is compact, then this weak* closure can be expressed as

$$\mathcal{A}_K^\infty = \bigcup_{B \times B \in \mathcal{P}_K} \mathcal{A}_{B^{\text{co}} \times B^{\text{co}}} = \bigcup_{B \times B \in \mathcal{P}_K} L^\infty(\Omega; B^{\text{co}}), \tag{4.3}$$

see [35, Remark 3.2 (b)] and [34, Theorem 1.1]. We point out the strong structural resemblance between (4.3) and the representation formula for $\mathcal{A}_K$ in (4.1). In both cases, it suffices to consider the nonlocal inclusions restricted to the maximal Cartesian sets of K, whose solutions trivially correspond to essentially bounded functions with a suitably restricted codomain.

Our guiding question in this section is the following:

Under what assumptions on K does the weak* closure of $\mathcal{A}_K$ preserve the structure of the nonlocal inclusion, that is, when is $\mathcal{A}_K^\infty = \mathcal{A}_L$ with a suitable $L \subset \mathbb{R}^m \times \mathbb{R}^m$?

According to [34, Theorem 1.1], structure preservation of $\mathcal{A}_K$ under weak* closures is always guaranteed when $m = 1$, and it was shown to hold under specific assumptions in the vectorial case. As the first step towards a complete picture for arbitrary dimensions, we start by identifying the Cartesian convexity of K as a necessary and sufficient condition for $\mathcal{A}_K$ being weakly* closed in $L^\infty(\Omega; \mathbb{R}^m)$, or in other words, for $\mathcal{A}_K^\infty = \mathcal{A}_K$.

Proposition 4.1. *Let $K \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric, diagonal, and compact. Then, $\mathcal{A}_K$ is L^∞-weakly* closed if and only if K is Cartesian convex.*

Proof. To prove the sufficiency, observe that the Cartesian convexity of K implies

$$K = \bigcup_{B \times B \in \mathcal{P}_K} B^{\text{co}} \times B^{\text{co}}$$

by Lemma 3.2(iv). Therefore, along with (4.3),

$$\mathcal{A}_K \subset \mathcal{A}_K^\infty = \bigcup_{B \times B \in \mathcal{P}_K} \mathcal{A}_{B^{\mathrm{co}} \times B^{\mathrm{co}}} \subset \mathcal{A}_{\bigcup_{B \times B \in \mathcal{P}_K} B^{\mathrm{co}} \times B^{\mathrm{co}}} = \mathcal{A}_K.$$

As this chain of inclusions turns into identities, the weak* closedness of $\mathcal{A}_K$ follows.

For the necessity, suppose that $\mathcal{A}_K^\infty = \mathcal{A}_K$. Then, again by (4.3), it holds for any $B \times B \in \mathcal{P}_K$ that $\mathcal{A}_{B^{\mathrm{co}} \times B^{\mathrm{co}}} \subset \mathcal{A}_K^\infty = \mathcal{A}_K$, so that $B^{\mathrm{co}} \times B^{\mathrm{co}} \subset K$ in view of (4.2). Consequently,

$$K = \bigcup_{B \times B \in \mathcal{P}_K} B^{\mathrm{co}} \times B^{\mathrm{co}},$$

which, according to Lemma 3.2, shows that K is Cartesian convex. □

The previous statement provides a natural candidate for the compact set $L \subset \mathbb{R}^m \times \mathbb{R}^m$ in the representation $\mathcal{A}_K^\infty = \mathcal{A}_L$ (if existent), namely,

$$L = K^{\times \mathrm{c}}. \tag{4.4}$$

Indeed, since L needs to be Cartesian convex by Proposition 4.1 and contains K because of $\mathcal{A}_K \subset \mathcal{A}_K^\infty = \mathcal{A}_L$ in view of (4.2), clearly, $L \supset K^{\times \mathrm{c}}$. On the other hand, it follows from (4.3) that

$$\mathcal{A}_L = \mathcal{A}_K^\infty = \bigcup_{B \times B \in \mathcal{P}_K} \mathcal{A}_{B^{\mathrm{co}} \times B^{\mathrm{co}}} \subset \mathcal{A}_{K_1^{\times \mathrm{c}}},$$

which implies $L \subset K_1^{\times \mathrm{c}} \subset K^{\times \mathrm{c}}$ in view of (4.2). Notice that $K_1^{\times \mathrm{c}}$ is, besides being symmetric and diagonal, also compact according to Lemma 3.6.

However, $\mathcal{A}_K^\infty = \mathcal{A}_{K^{\times \mathrm{c}}}$ cannot be true for all K; we have implicitly encountered a limitation already in the previous reasoning, which requires

$$K^{\times \mathrm{c}} = K_1^{\times \mathrm{c}}.$$

Recalling that the Cartesian convex hull of K is called basic in the sense of Definition 3.8 if $\mathcal{P}_{K^{\times \mathrm{c}}} \subset \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_K\}$ allows us to deduce the following characterization result.

Proposition 4.2. *Let $K \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric, diagonal, and compact. Then,*

$$\mathcal{A}_K^\infty = \mathcal{A}_{K^{\times \mathrm{c}}} \qquad \textit{if and only if} \qquad K \textit{ has a basic Cartesian convexification.}$$

The proof is a direct consequence of the next lemma.

Lemma 4.3. *Let $K, L \subset \mathbb{R}^m \times \mathbb{R}^m$ be symmetric, diagonal, and compact. Then, $\mathcal{A}_K^\infty = \mathcal{A}_L$ if and only if*

$$L = K^{\times \mathrm{c}} \quad \textit{and} \quad \mathcal{P}_L \subset \{B^{\mathrm{co}} \times B^{\mathrm{co}} : B \times B \in \mathcal{P}_K\}. \tag{4.5}$$

Proof. If $\mathcal{A}_K^\infty = \mathcal{A}_L$, then $L = K^{\times \mathrm{c}}$ by (4.4). To show also the second part of (4.5), let $D \subset \mathbb{R}^m$ with $D \times D \in \mathcal{P}_L$. Moreover, take $D' = \{d_1, d_2, \ldots\}$ to be a countable dense subset of D, e.g. $D' = D \cap \mathbb{Q}^m$, and set $D_i := \{d_1, \ldots, d_i\}$ for $i \in \mathbb{N}$; note that these sets are nested, that is, $D_i \subset D_{i+1}$ for $i \in \mathbb{N}$, and satisfy $D' = \bigcup_{i \in \mathbb{N}} D_i$.

If we consider for each $i \in \mathbb{N}$ a simple function $u_i : \Omega \to D_i$ such that each preimage $u_i^{-1}(d_j)$ for $j = 1, \ldots, i$ has positive Lebesgue measure, then, by construction, $u_i \in \mathcal{A}_L = \mathcal{A}_K^\infty$ and thus, by (4.3), $u_i \in L^\infty(\Omega; B_i^{\mathrm{co}})$ for some $B_i \times B_i \in \mathcal{P}_K$.

Applying Lemma A.1 to the sequence of compact sets $(B_i)_i$ gives a compact set $B \subset \mathbb{R}^m$ with $B \times B \subset K$ such that

$$\sup_{\xi \in D'} \text{dist}(\xi, B^{\text{co}}) = \lim_{i \to \infty} \sup_{\xi \in D_i} \text{dist}(\xi, B^{\text{co}}) \leq \liminf_{i \to \infty} d_H(B_i^{\text{co}}, B^{\text{co}}) = 0.$$

Hence, $D' \subset B^{\text{co}}$, and taking the closure implies $D \subset B^{\text{co}}$ and

$$D \times D \subset B^{\text{co}} \times B^{\text{co}} \subset K^{\times \text{c}}.$$

After enlarging B, we may assume that $B \times B \in \mathcal{P}_K$, and it follows due to the maximality of $D \times D$ in $L = K^{\times \text{c}}$ that $D = B^{\text{co}}$.

Now, suppose that (4.5) holds. Since $L = K^{\times \text{c}}$ is Cartesian convex, we infer from Proposition 4.1 that $\mathcal{A}_L^\infty = \mathcal{A}_L$. It follows then together with (4.1), (4.3), and $K \subset K^{\times \text{c}} = L$ that

$$\mathcal{A}_L = \bigcup_{D \times D \in \mathcal{P}_L} L^\infty(\Omega; D) \subset \bigcup_{B \times B \in \mathcal{P}_K} L^\infty(\Omega; B^{\text{co}}) = \mathcal{A}_K^\infty \subset \mathcal{A}_L^\infty = \mathcal{A}_L,$$

which yields $\mathcal{A}_K^\infty = \mathcal{A}_L$. $\square$

Observe that the previous findings are consistent with [34, Theorem 1.1], which is actually a special case of Proposition 4.2. Suppose $K \subset \mathbb{R}^m \times \mathbb{R}^m$ is symmetric, diagonal, and compact such that K^{sc} is compact and satisfies

$$\widehat{K^{\text{sc}}} = \bigcup_{(\alpha, \beta) \in K} [\alpha, \beta] \times [\alpha, \beta] \quad \text{and} \quad \mathcal{P}_{\widehat{K^{\text{sc}}}} = \mathcal{P}_{K^{\text{sc}}} \subset \{[\alpha, \beta] \times [\alpha, \beta] : \{\alpha, \beta\} \times \{\alpha, \beta\} \in K\}. \tag{4.6}$$

Under these assumptions, one can mimic the one-dimensional argument of Remark 3.3 (a) to derive the inclusion $\widehat{K^{\text{sc}}} \subset K_1^{\times \text{c}}$. Together with $K^{\times \text{c}} \subset \widehat{K^{\text{sc}}}$, it follows that

$$\widehat{K^{\text{sc}}} = K_1^{\times \text{c}} = K^{\times \text{c}}. \tag{4.7}$$

Since the second part of (4.6) in combination with (4.7) implies that the Cartesian convex hull of K is basic, we can apply Proposition 4.2 to infer $\mathcal{A}_K^\infty = \mathcal{A}_{K_1^{\times \text{c}}}$, and hence,

$$\mathcal{A}_K^\infty = \mathcal{A}_{\widehat{K^{\text{sc}}}} = \mathcal{A}_{K^{\text{sc}}},$$

as stated in [34, Theorem 1.1].

We conclude our discussion of nonlocal inclusions with an alternative representation of $\mathcal{A}_K^\infty$ via barycenters of Young measures. It serves as a useful technical tool in the final section of this paper on L^p-approximation.

Remark 4.4 (*Young Measure Representation of $\mathcal{A}_K^\infty$*)**.** For any $K \subset \mathbb{R}^m \times \mathbb{R}^m$ symmetric and compact, let

$$\mathcal{V}_K := \{\nu \in L_w^\infty(\Omega; Pr(\mathbb{R}^m)) : \text{supp}(\nu_x \otimes \nu_y) \subset K \text{ for a.e. } (x, y) \in \Omega \times \Omega\}. \tag{4.8}$$

We claim that

$$\mathcal{A}_K^\infty = \{u \in L^\infty(\Omega; \mathbb{R}^m) : u = [\nu], \nu \in \mathcal{V}_K\}, \tag{4.9}$$

or in other words,

$$\nu \in \mathcal{V}_K \qquad \text{if and only if} \qquad [\nu] \in \mathcal{A}_K^\infty. \tag{4.10}$$

This characterization of $\mathcal{A}_K^\infty$ is essentially a reformulation of the results in [35, Theorems 3.3 and 3.1]. A sketch of the corner points of the derivation is included below for the reader's convenience; for further details, see [35, Section 3]. As for notation, we associate with every $u \in L^\infty(\Omega; \mathbb{R}^m)$, the nonlocal vector field

$$v_u(x, y) := (u(x), u(y)) \qquad \text{for } (x, y) \in \Omega \times \Omega. \tag{4.11}$$

The principal identity behind (4.9) and (4.10) is

$$\mathcal{Y}_K^\infty = \mathcal{Y}_K \tag{4.12}$$

from [35, Theorem 3.3], where

$$\mathcal{Y}_K^\infty = \{\Lambda \in L_w^\infty(\Omega \times \Omega; \mathcal{P}r(\mathbb{R}^m \times \mathbb{R}^m)) : v_{u_j} \xrightarrow{\mathrm{YM}} \Lambda, \text{ with } (u_j)_j \subset \mathcal{A}_K\}$$

and

$$\mathcal{Y}_K = \{\Lambda \in L_w^\infty(\Omega \times \Omega; \mathcal{P}r(\mathbb{R}^m \times \mathbb{R}^m)) : \Lambda_{(x,y)} = \nu_x \otimes \nu_y \text{ with } \nu \in L_w^\infty(\Omega; \mathcal{P}r(\mathbb{R}^m)),$$
$$\mathrm{supp}(\nu_x \otimes \nu_y) \subset K \text{ a.e. } (x,y) \in \Omega \times \Omega\}$$
$$= \{\Lambda \in L_w^\infty(\Omega \times \Omega; \mathcal{P}r(\mathbb{R}^m \times \mathbb{R}^m)) : \Lambda_{(x,y)} = \nu_x \otimes \nu_y \text{ with } \nu \in \mathcal{V}_K\}.$$

The proof of (4.12) relies on characterization results of Young measures generated by sequences of both exact and approximate solutions to the nonlocal inclusion problem, together with an approximating argument for elements in $\mathcal{Y}_K$ by suitable inhomogeneous convex combinations of Dirac masses.

By taking barycenters, we obtain with the same reasoning as in the proof of [35, Theorem 3.1] that

$$\mathcal{A}_K^\infty = \{u \in L^\infty(\Omega; \mathbb{R}^m) : v_u = [\Lambda], \Lambda \in \mathcal{Y}_K^\infty\} = \{u \in L^\infty(\Omega; \mathbb{R}^m) : v_u = [\Lambda], \Lambda \in \mathcal{Y}_K\}$$
$$= \{u \in L^\infty(\Omega; \mathbb{R}^m) : u = [\nu], \nu \in \mathcal{V}_K\}, \quad \triangle$$

as stated (4.9).

5. Lower semicontinuity and relaxation of nonlocal supremals

In this section, we consider nonlocal supremal functionals of the form J_W as defined in (1.1), with a lower semicontinuous and coercive supremand $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. Since $J_W = J_{\widehat{W}}$, that is,

$$\operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} W(u(x), u(y)) = \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} \widehat{W}(u(x), u(y)) \tag{5.1}$$

for every $u \in L^\infty(\Omega; \mathbb{R}^m)$ (see [34, (7.3)]), we may assume without loss of generality that W is also symmetric and diagonal.

With Proposition 4.1 from the previous section at hand, we are now in the position to prove Theorem 1.1, which identifies Cartesian level convexity of the supremand as necessary and sufficient for the weak* lower semicontinuity of nonlocal L^∞-functionals. Considering that the notions of Cartesian level convexity and separate level convexity coincide in the scalar case (cf. Remark 3.3 (a)), this generalizes the findings of [34, Theorem 1.3 (i)] to arbitrary dimensions.

Proof of Theorem 1.1. By [34, Proposition 7.1], the L^∞-weak* lower semicontinuity of J_W is equivalent to the L^∞-weak* closedness of $\mathcal{A}_{L_c(W)}$ for any $c \in \mathbb{R}$. Since all sublevel sets of W are symmetric, diagonal, and compact in light of the assumptions on W, the stated equivalence is an immediate consequence of Proposition 4.1. $\square$

When J_W fails to be weakly* lower semicontinuous, finding a good representation of its weakly* lower semicontinuous envelope (see (1.3)) is a central question in relaxation theory. Our next auxiliary result establishes the connection between the relaxation of nonlocal supremal functionals and the L^∞-weak* closure of their supremands' sublevel sets; it can be viewed as an extension of [34, Proposition 7.1].

Lemma 5.1.   *Let $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be symmetric, diagonal, lower semicontinuous, and coercive, and let $I : L^\infty(\Omega; \mathbb{R}^m) \to \mathbb{R}$ be a functional. Then, the following identities are equivalent:*

(i) $I = J_W^{\mathrm{rlx}}$;
(ii) $L_c(I) = \mathcal{A}_{L_c(W)}^\infty$ for all $c \in \mathbb{R}$;
(iii) $I(u) = \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)}^\infty\}$.

Proof. The proof is straightforward, nevertheless, we detail the arguments here for the reader's convenience.

$(ii) \Rightarrow (i)$ To prove that I is a lower bound on the relaxation of J_W, consider $(u_j)_j \subset L^\infty(\Omega; \mathbb{R}^m)$ such that $u_j \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega; \mathbb{R}^m)$, assuming without loss of generality that

$$\liminf_{j\to\infty} J_W(u_j) = \lim_{j\to\infty} J_W(u_j) =: c < \infty.$$

Fixing $\varepsilon > 0$, we find $(u_j)_j \subset \mathcal{A}_{L_{c+\varepsilon}(W)}$ for j sufficiently large, and thus, $u \in \mathcal{A}_{L_{c+\varepsilon}(W)}^\infty$, so that $I(u) \le c + \varepsilon = \liminf_{j\to\infty} J_W(u_j) + \varepsilon$. Since $(u_j)_j$ is an arbitrary weakly* converging sequence with limit u, taking the limit $\varepsilon \to 0$ gives $I(u) \le J_W^{\mathrm{rlx}}(u)$.

For the reverse inequality, we observe that any $u \in L^\infty(\Omega; \mathbb{R}^m)$ satisfies $u \in L_{I(u)}(I) = \mathcal{A}_{L_{I(u)}(W)}^\infty$ and can therefore be approximated in the sense of L^∞-weak* convergence by a sequence $(u_j)_j \subset \mathcal{A}_{L_{I(u)}(W)}$, so that

$$J_W^{\mathrm{rlx}}(u) \le \limsup_{j\to\infty} J_W(u_j) = \limsup_{j\to\infty} \operatorname*{ess\,sup}_{(x,y)\in\Omega\times\Omega} W(u_j(x), u_j(y)) \le I(u).$$

$(i) \Rightarrow (ii)$ We show first that any $u \in L_c(I) = L_c(J_W^{\mathrm{rlx}})$ for $c \in \mathbb{R}$ is the weak* limit of a sequence in $\mathcal{A}_{L_c(W)}$. To that end, let $\varepsilon > 0$ and take a recovery sequence $(u_j)_j \subset L^\infty(\Omega; \mathbb{R}^m)$ for u. Then, $u_j \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega; \mathbb{R}^m)$ and $\lim_{j\to\infty} J_W(u_j) = J_W^{\mathrm{rlx}}(u) \le c$. The latter implies $u_j \in \mathcal{A}_{L_{c+\varepsilon}(W)}$ for all $j \in \mathbb{N}$ sufficiently large. Hence, $u \in \mathcal{A}_{L_{c+\varepsilon}(W)}^\infty$, and since $\varepsilon > 0$ was arbitrary, $u \in \mathcal{A}_{L_c(W)}^\infty$ by Lemma A.3.

On the other hand, if $u \in \mathcal{A}_{L_c(W)}^\infty$ with $u_j \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega; \mathbb{R}^m)$ and $(u_j)_j \subset \mathcal{A}_{L_c(W)}$, then

$$c \ge \liminf_{j\to\infty} J_W(u_j) \ge J_W^{\mathrm{rlx}}(u) = I(u),$$

meaning $u \in L_c(I)$.

$(ii) \Leftrightarrow (iii)$ This follows from (2.4) and (2.5) while taking Lemma A.2 into account. $\square$

By combining the statement of the previous lemma with the characterization of weak* closures of nonlocal inclusions in (4.3), we infer that

$$J_W^{\mathrm{rlx}}(u) = \inf\{c \in \mathbb{R} : u \in L^\infty(\Omega; B^{\mathrm{co}}) \text{ with } B \times B \in \mathcal{P}_{L_c(W)}\}$$

for $u \in L^\infty(\Omega; \mathbb{R}^m)$. Even if this provides a new and simpler representation of the relaxation of J_W, it is not obvious from this formula whether (and – if not in general – under what assumptions) J_W^{rlx} corresponds again to a supremal functional of the form (1.1). We give a characterization result about when J_W^{rlx} is structure-preserving in the next section.

5.1. Structure preservation under relaxation

As a consequence of Theorem 1.1, we can immediately deduce a lower bound on the relaxation of J_W as in (1.1) in the form of another nonlocal supremal of the same type, precisely,

$$J_W^{\mathrm{rlx}} \ge J_{\widehat{W}^{\times \mathrm{lc}}}, \tag{5.2}$$

where $\widehat{W}^{\times \mathrm{lc}}$ is the Cartesian level convex envelope of $\widehat{W}$, see Definition 3.10.

Our second main result of the paper, formulated in Theorem 1.2, characterizes the conditions on W under which equality holds in (5.2). It is a direct implication of the following proposition, which again relies on the characterization of weak* closures of nonlocal inclusions from Proposition 4.2.

Proposition 5.2. *Let $W : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be symmetric, diagonal, lower semicontinuous, and coercive. Then, the following two statements are equivalent:*

(i) all sublevel sets of W admit a basic Cartesian convexification;
(ii) there exists a symmetric, diagonal, lower semicontinuous, and coercive function $G : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ such that $J_W^{\mathrm{rlx}} = J_G$.

If these conditions are fulfilled, then $J_W^{\mathrm{rlx}} = J_{W^{\times \mathrm{lc}}}$.

Proof. $(i) \Rightarrow (ii)$ The Cartesian level convexification $W^{\times \mathrm{lc}}$ is symmetric, diagonal, lower semicontinuous, and coercive, and satisfies
$$L_c(W^{\times \mathrm{lc}}) = L_c(W)^{\times \mathrm{c}} \quad \text{for all } c \in \mathbb{R},$$
according to Lemma 3.13. Moreover, we infer from Lemma 5.1 and Proposition 4.2 (see also (2.4)) that

$$J_W^{\mathrm{rlx}}(u) = \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)}^{\infty}\} = \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)^{\times \mathrm{c}}}\}$$
$$= \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W^{\times \mathrm{lc}})}\} = \inf\{c \in \mathbb{R} : u \in L_c(J_{W^{\times \mathrm{lc}}})\} = J_{W^{\times \mathrm{lc}}}(u)$$

for $u \in L^{\infty}(\Omega; \mathbb{R}^m)$. Hence, setting $G = W^{\times \mathrm{lc}}$ proves (ii).

$(ii) \Rightarrow (i)$ As (ii) holds, we can use once again Lemma 5.1 (and (2.4)) to obtain

$$J_G(u) = J_W^{\mathrm{rlx}}(u) = \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)}^{\infty}\}$$

and consequently,

$$\mathcal{A}_{L_c(W)}^{\infty} = L_c(J_G) = \mathcal{A}_{L_c(G)} \quad \text{for every } c \in \mathbb{R}$$

by (2.5) under consideration of Lemma A.3. Now, the claim follows immediately from Lemma 4.3. $\square$

Summing up, we have seen that, in general, J_W^{rlx} is not a supremal functional of the form (1.1). The next section discusses different (counter)examples of representation formulas in cases when W are specific multi-well supremands.

5.2. Relaxation of nonlocal supremals in multi-well form

The intention of this section is to illustrate the previous results, using the example of L^{∞}-functionals

$$J_W(u) := \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} W(u(x), u(y)) \quad \text{for } u \in L^{\infty}(\Omega; \mathbb{R}^m)$$

with multi-well supremands of the form
$$W = \min_{i=1,\dots,N} W_i,$$

where $W_i : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is symmetric, diagonal, lower semicontinuous and coercive for $i = 1, \dots, N$ with $N \in \mathbb{N}$.

Below, we will often specify this setting further by choosing W_i as single-well functions

$$W_i(\xi, \zeta) = \operatorname{dist}\big((\xi, \zeta), A_i \times A_i\big) \quad \text{for } (\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m \tag{5.3}$$

with compact sets $A_i \subset \mathbb{R}^m$ such that (2.13) holds. Moreover, for $(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m$,

$$W(\xi, \zeta) = \operatorname{dist}\big((\xi, \zeta), K\big) \quad \text{with } K := \bigcup_{i=1}^{N} A_i \times A_i. \tag{5.4}$$

Observe for the sublevel sets that

$$L_c(W) = \bigcup_{i=1}^{N} (A_i)_c \times (A_i)_c = \bigcup_{i=1}^{N} L_c(W_i)$$

with $c \in \mathbb{R}$ and $A_c := A + B_c(0)$ for $A \subset \mathbb{R}^m$.

Before we focus on relaxation formulas with and without structure preservation (see Examples 5.5 and 5.6), let us provide the following basic representation for general multi-well L^∞-functionals.

Lemma 5.3. *Let* $W_i : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ *for* $i = 1, \ldots, N$ *with* $N \in \mathbb{N}$ *be symmetric, diagonal, lower semicontinuous, and coercive, and suppose that* $W = \min_{i=1,\ldots,N} W_i$ *satisfies*

$$\mathcal{P}_{L_c(W)} \subset \bigcup_{i=1}^{N} \mathcal{P}_{L_c(W_i)} \quad \text{for all } c \in \mathbb{R}. \tag{5.5}$$

Then,

$$J_W = \min_{i=1,\ldots,N} J_{W_i}. \tag{5.6}$$

Proof. Since the sublevel sets of W are given by $L_c(W) = \bigcup_{i=1}^{N} L_c(W_i)$ for $c \in \mathbb{R}$, it is immediate to check that the properties of the functions W_i carry directly over to W, that is, also W is symmetric, diagonal, lower semicontinuous, and coercive. Considering (2.4), (4.1), and (5.5), this can be seen like this: For any $u \in L^\infty(\Omega; \mathbb{R}^m)$,

$$\begin{aligned}
J_W(u) &= \inf\big\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)}\big\} = \inf\big\{c \in \mathbb{R} : u \in \bigcup\nolimits_{B \times B \in \mathcal{P}_{L_c(W)}} L^\infty(\Omega; B)\big\} \\
&\geq \inf\{c \in \mathbb{R} : u \in \bigcup\nolimits_{B \times B \in \mathcal{P}_{L_c(W_i)}} L^\infty(\Omega; B) \text{ for some } i \in \{1, \ldots, N\}\} \\
&= \min_{i=1,\ldots,N} \inf\{c \in \mathbb{R} : u \in \bigcup\nolimits_{B \times B \in \mathcal{P}_{L_c(W_i)}} L^\infty(\Omega; B)\} \\
&= \min_{i=1,\ldots,N} \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W_i)}\} = \min_{i=1,\ldots,N} J_{W_i}(u).
\end{aligned}$$

The reverse inequality is trivially satisfied due to $W \leq W_i$ for all i. $\quad\square$

Notice that the assumption (5.5) is necessary for proving (5.6) for $N > 2$, as we show below for three-well supremands W. In particular, this highlights what implications the intrinsic non-uniqueness in the representation of W as minimum of the functions W_i has for the representation of J_W and illustrates the role of hidden squares of the sublevel sets of W.

Example 5.4. Let W and W_i for $i = 1, \ldots, N$ be as in (5.4) and (5.3).

(a) In case W is a double-well supremand, which corresponds to $N = 2$, one always has $J_W = \min_{i=1,2} J_{W_i}$. Indeed, the condition (5.5) is satisfied here, as Conclusion 2.7 (b) implies

$$\mathcal{P}_{L_c(W)} \subset \{(A_1)_c \times (A_1)_c\} \cup \{(A_2)_c \times (A_2)_c\} = \mathcal{P}_{L_c(W_1)} \cup \mathcal{P}_{L_c(W_2)}$$

for all $c \in \mathbb{R}$.

(b) If $N = 3$, meaning that W is a three-well supremand, it is in general not true that $J_W = \min_{i=1,2,3} J_{W_i}$. This observation is tied to the occurrence of hidden maximal Cartesian squares, which can

affect the structure of nonlocal supremal functionals. Taking for instance compact sets $A_1, A_2, A_3 \subset \mathbb{R}^m$ as in Example 2.3, then $M \times M \in \mathcal{P}_K = \mathcal{P}_{L_0(W)}$ with M defined in (2.11), but $M \times M \notin \mathcal{P}_{L_0(W_i)} = \{A_i \times A_i\}$ for any $i = 1, 2, 3$.

However, we will prove that it is possible to express W as the minimum of four functions so that the condition (5.5) is satisfied. To this aim, consider the function $W_4 : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ defined for $(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m$ by

$$W_4(\xi, \zeta) = \inf\{c \in \mathbb{R} : (\xi, \zeta) \in M^c \times M^c\} \tag{5.7}$$

with

$$M^c := \bigcup_{i,j=1,2,3, i \neq j} (A_i)_c \cap (A_j)_c \quad \text{for all } c \in \mathbb{R}. \tag{5.8}$$

Then, $L_c(W_4) = M^c \times M^c \subset L_c(W) = \bigcup_{i=1}^3 L_c(W_i)$ in view of (2.5) implies $W \leq W_4$ and thus,

$$W = \min_{i=1,\ldots,4} W_i. \tag{5.9}$$

By Lemma 2.6, it is straightforward to verify that (5.5) is indeed fulfilled for this representation of W, and we conclude from Lemma 5.3 that

$$J_W = \min_{i=1,2,3,4} J_{W_i}. \quad \triangle$$

In view of Lemma 5.3, it is not hard to see that the relaxation of J_W with $W = \min_{i=1,\ldots,N} W_i$ satisfying (5.5) can be obtained via minimization over the relaxed supremals associated to the more basic functions W_i, that is,

$$J_W^{\mathrm{rlx}} = \min_{i=1,\ldots,N} J_{W_i}^{\mathrm{rlx}}. \tag{5.10}$$

Indeed, as a consequence of (5.6), we infer for the relaxation of J_W that

$$J_{W_j}^{\mathrm{rlx}} \geq J_W^{\mathrm{rlx}} \geq \left(\min_{i=1,\ldots,N} J_{W_i}^{\mathrm{rlx}}\right)^{\mathrm{rlx}}$$

for all $j \in \{1, \ldots, N\}$. The identity (5.10) follows then by taking the minimum over j, along with the observation that the minimum of finitely many weakly* lower semicontinuous functionals is again weakly* lower semicontinuous.

Without further hypotheses on W_i, it is not possible to deduce $J_{W_i}^{\mathrm{rlx}} = J_{W_i^{\times \mathrm{lc}}}$, as Proposition 5.2 indicates. If W_i is a single-well supremands like in (5.3), though, then J_{W_i} is actually a classical L^∞-functional, that is,

$$J_{W_i}(u) = \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} \mathrm{dist}((u(x), u(y)), A_i \times A_i) = \operatorname*{ess\,sup}_{x \in \Omega} \mathrm{dist}(u(x), A_i),$$

cf. (2.1). The relaxation of J_{W_i} can therefore be determined via the classical theory of L^∞-functionals (see [44, Theorem 2.5]), which gives

$$J_{W_i}^{\mathrm{rlx}}(u) = \operatorname*{ess\,sup}_{x \in \Omega} \mathrm{dist}(u(x), A_i^{\mathrm{co}})$$

for $u \in L^\infty(\Omega; \mathbb{R}^m)$, or equivalently,

$$J_{W_i}^{\mathrm{rlx}} = J_{W_i^{\mathrm{co}}} = J_{W_i^{\times \mathrm{lc}}};$$

here $W_i^{\times \mathrm{lc}}(\xi, \zeta) = W_i^{\mathrm{co}}(\xi, \zeta) = \mathrm{dist}((\xi, \zeta), A^{\mathrm{co}} \times A^{\mathrm{co}})$ for $(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m$ and W^{co} denotes the convex envelope of W.

With these preparations, we now present two qualitatively different examples of relaxation results for nonlocal supremals of three-well type, which build on the insights from Examples 3.5 and 5.4. Even though, the relaxed functionals can both be calculated via the same formula through minimization over four nonlocal supremals (see (5.11) below), one case is structure-preserving, while the other is not.

Example 5.5 (*Structure Preservation During Relaxation*). Consider W and W_i for $i = 1, 2, 3$ as (5.3) and (5.4) with convex sets $A_1, A_2, A_3 \subset \mathbb{R}^m$ with pairwise nonempty intersections. Then not only do the three sets A_1, A_2, A_3 fit into the framework of Example 5.4 (b), but more generally, also $(A_i)_c = A_i + B_c(0)$ with $i = 1, 2, 3$ for any $c \in \mathbb{R}$. Hence, all sublevel sets of W have a basic Cartesian convexification and

$$L_c(W)^{\times c} = \bigcup_{i=1}^{3} [(A_i)_c \times (A_i)_c] \cup [(M^c)^{co} \times (M^c)^{co}] = \bigcup_{i=1}^{4} L_c(W_i)^{\times c}$$

for any $c \in \mathbb{R}$ with M^c and W_4 as in (5.8) and (5.7), respectively. This allows us to conclude with Proposition 5.2 that

$$J_W^{rlx} = J_{W^{\times lc}},$$

where the Cartesian level convex envelope of W can be determined via

$$W^{\times lc} = \min_{i=1,2,3,4} W_i^{\times lc} = \min_{i=1,2,3,4} W_i^{co}. \qquad \triangle$$

Example 5.6 (*Loss of Nonlocal Supremal Structure During Relaxation*). Let W and W_i for $i = 1, 2, 3$ as (5.4) and (5.3) with three sets $A_1, A_2, A_3 \subset \mathbb{R}^m$ as in Example 3.5 (b). Then $L_0(W) = K$ does not admit a basic Cartesian convexification, since

$$\mathcal{P}_{L_0(W)^{\times c}} = \mathcal{P}_{K^{\times c}} = \bigcup_{i=1}^{3} \{A_i^{co} \times A_i^{co}\} \cup \{M^{co} \times M^{co}\}$$

but $M \times M \notin \mathcal{P}_{L_0(W)}$. As a consequence of Proposition 5.2,

$$J_W^{rlx} \neq J_{W^{\times lc}}. \qquad \triangle$$

We point out that in both examples above, J_W^{rlx} can be expressed through minimization over the relaxations of the more basic supremal functionals, namely,

$$J_W^{rlx} = \min_{i=1,2,3,4} J_{W_i}^{rlx} = \min_{i=1,2,3,4} J_{W_i^{co}} = \min_{i=1,2,3,4} J_{W_i^{\times lc}}, \tag{5.11}$$

where W_4 is defined in (5.7); this is a consequence of (5.10) and (5.9). And still, only Example 5.5 allows us to rewrite J_W^{rlx} as a nonlocal supremal functional.

6. L^p-Approximation

This section provides the proof of the Γ-convergence result formulated in Theorem 1.3. Some of the arguments are inspired by [19, Theorem 3.1], where the authors show a power-law approximation result connecting single integrals and classical (local) supremal functionals.

Proof of Theorem 1.3. We subdivide the proof into three natural steps, starting with the relative weak sequential compactness in $L^q(\Omega; \mathbb{R}^m)$ of sequences of uniformly bounded energy and followed by the lower and upper bounds for the Γ-limits of $(I_W^p)_{p \geq q}$, see Definition 2.2. Both bounds exploit the classical L^p-approximation of the L^∞-norm as well as the relaxation result established in Section 5.

In the following, let $(p_j)_j \subset [q, \infty)$ be a sequence that converges indefinitely to ∞, i.e., $p_j \to \infty$ as $j \to \infty$. Besides, suppose that $c_2 = 0$ in the growth condition (1.6); otherwise, consider the translated double-integrand $W + c_2$ in place of W.

Step 1: Compactness. Let $(u_j)_j \subset L^q(\Omega; \mathbb{R}^m)$ with

$$\sup_{j \in \mathbb{N}} I_W^{p_j}(u_j) < \infty. \tag{6.1}$$

We show that $(u_j)_j$ is bounded in $L^q(\Omega; \mathbb{R}^m)$ which yields the existence of a subsequence of $(u_j)_j$ that converges weakly in $L^q(\Omega; \mathbb{R}^m)$.

The growth condition on W from below (see (1.6) with $c_2 = 0$) along with Hölder's inequality and $p_j \geq q$, gives for each such j that

$$I_W^{p_j}(u_j) = \|W(v_{u_j})\|_{L^{p_j}(\Omega \times \Omega)} \geq \mathcal{L}^n(\Omega)^{2(\frac{1}{p_j} - \frac{1}{q})} \|W(v_{u_j})\|_{L^q(\Omega \times \Omega)} \tag{6.2}$$
$$\geq c_1 \mathcal{L}^n(\Omega)^{2(\frac{1}{p_j} - \frac{1}{q})} \|v_{u_j}\|_{L^q(\Omega \times \Omega; \mathbb{R}^m \times \mathbb{R}^m)} \geq c \mathcal{L}^n(\Omega)^{\frac{2}{p_j}} \|u_j\|_{L^q(\Omega; \mathbb{R}^m)},$$

where v_{u_j} are the nonlocal fields as defined in (4.11) and $c > 0$ is a constant independent of j.

In light of (6.1) and $\lim_{j \to \infty} \mathcal{L}^n(\Omega)^{1/p_j} = 1$, this shows that $(u_j)_j$ is uniformly bounded in $L^q(\Omega; \mathbb{R}^m)$, which implies immediately the existence of a weakly convergent subsequence of $(u_j)_j$ in $L^q(\Omega; \mathbb{R}^m)$.

Step 2: Lower bound. Let $(u_j)_j \subset L^q(\Omega; \mathbb{R}^m)$ and $u \in L^q(\Omega; \mathbb{R}^m)$ such that $u_j \rightharpoonup u$ in $L^q(\Omega; \mathbb{R}^m)$. To show that

$$\liminf_{j \to \infty} I_W^{p_j}(u_j) \geq I_W^\infty(u),$$

we may assume without loss of generality that

$$\lim_{j \to \infty} I_W^{p_j}(u_j) = \liminf_{j \to \infty} I_W^{p_j}(u_j) < \infty,$$

and thus, $u_j \in L^{p_j}(\Omega; \mathbb{R}^m)$ for all $j \in \mathbb{N}$. As a consequence of the estimate (6.2), one has for any $r \geq q > 1$ that

$$\lim_{j \to \infty} I_W^{p_j}(u_j) \geq \mathcal{L}^n(\Omega)^{-\frac{2}{r}} \liminf_{j \to \infty} I_W^r(u_j), \tag{6.3}$$

and $(u_j)_j$ can be regarded as uniformly bounded in $L^r(\Omega; \mathbb{R}^m)$. The latter allows us, in view of Proposition 2.1(i), to extract a subsequence (not relabeled) generating a Young measure $\nu \in L_w^\infty(\Omega; \mathcal{P}r(\mathbb{R}^m))$ with $u = [\nu]$, that is,

$$u_j \xrightarrow{\text{YM}} \nu \quad \text{as } j \to \infty.$$

Then, by [41, Proposition 2.3], also the associated nonlocal fields $(v_{u_j})_j$ generate a Young measure, namely,

$$v_{u_j} \xrightarrow{\text{YM}} \Lambda \quad \text{as } j \to \infty \qquad \text{with } \Lambda_{(x,y)} = \nu_x \otimes \nu_y \text{ for a.e. } (x,y) \in \Omega \times \Omega,$$

and Proposition 2.1(ii) implies

$$\liminf_{j \to \infty} I_W^r(u_j) = \liminf_{j \to \infty} \left(\int_\Omega \int_\Omega W^r(u_j(x), u_j(y)) \, dx \, dy \right)^{\frac{1}{r}}$$
$$\geq \left(\int_\Omega \int_\Omega \int_{\mathbb{R}^m \times \mathbb{R}^m} W^r(\xi, \zeta) \, d(\nu_x \otimes \nu_y)(\xi, \zeta) \, dx \, dy \right)^{\frac{1}{r}}.$$

For any $r > s > 1$, we can exploit the convexity of the map $t \mapsto t^{\frac{r}{s}}$ to obtain

$$\liminf_{j \to \infty} I_W^r(u_j) \geq \left(\int_\Omega \int_\Omega \left(\int_{\mathbb{R}^m \times \mathbb{R}^m} W^s(\xi, \zeta) \, d(\nu_x \otimes \nu_y)(\xi, \zeta) \right)^{\frac{r}{s}} dx \, dy \right)^{\frac{1}{r}}.$$

Now, passing successively to the limits $r \to \infty$ and $s \to \infty$ leads via iterative classical approximation of the L^∞-norm by (2.6) (see also (2.7)). to

$$\liminf_{r \to \infty} \liminf_{j \to \infty} I_W^r(u_j) \geq \liminf_{s \to \infty} \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} \left(\int_{\mathbb{R}^m \times \mathbb{R}^m} W^s(\xi, \zeta) \, d(\nu_x \otimes \nu_y)(\xi, \zeta) \right)^{\frac{1}{s}}$$
$$\geq \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} [(\nu_x \otimes \nu_y)\text{-} \operatorname*{ess\,sup}_{(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m} W(\xi, \zeta)].$$

To sum up, we conclude together with (6.3) and $\lim_{r \to \infty} \mathcal{L}^n(\Omega)^{\frac{2}{r}} = 1$ that

$$\lim_{j \to \infty} I_W^{p_j}(u_j) \geq \liminf_{r \to \infty} \liminf_{j \to \infty} I_W^r(u_j) \geq \operatorname*{ess\,sup}_{(x,y) \in \Omega \times \Omega} f_{W,\nu}(x,y) = \|f_{W,\nu}\|_{L^\infty(\Omega \times \Omega)},$$

where $f_{W,\nu}$ is defined as

$$f_{W,\nu}(x,y) := (\nu_x \otimes \nu_y)\text{-} \operatorname*{ess\,sup}_{(\xi, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m} W(\xi, \zeta)$$

for $(x,y) \in \Omega \times \Omega$. The proof of the liminf-inequality will be complete, once it is shown that

$$\|f_{W,\nu}\|_{L^\infty(\Omega \times \Omega)} = I_W^\infty(u). \tag{6.4}$$

To this end, observe that the lower semicontinuity of W implies

$$f_{W,\nu}(x,y) = \sup_{(\xi, \zeta) \in \operatorname{supp}(\nu_x \otimes \nu_y)} W(\xi, \zeta)$$

for $(x,y) \in \Omega \times \Omega$ (see (2.3)), and therefore,

$$\|f_{W,\nu}\|_{L^\infty(\Omega \times \Omega)} = \inf\{c \in \mathbb{R} : f_{W,\nu}(x,y) \leq c \text{ for a.e. } (x,y) \in \Omega \times \Omega\}$$
$$= \inf\{c \in \mathbb{R} : \operatorname{supp}(\nu_x \otimes \nu_y) \subset L_c(W) \text{ for a.e. } (x,y) \in \Omega \times \Omega\}$$
$$= \inf\{c \in \mathbb{R} : \nu \in \mathcal{V}_{L_c(W)}\},$$

with $\mathcal{V}_{L_c(W)}$ as introduced in (4.8). In combination with the equivalence from (4.10), stating that

$$u = [\nu] \in \mathcal{A}_{L_c(W)}^\infty \qquad \text{if and only if} \qquad \nu \in \mathcal{V}_{L_c(W)}$$

for any $c \in \mathbb{R}$, it follows that

$$\|f_{W,\nu}\|_{L^\infty(\Omega \times \Omega)} = \inf\{c \in \mathbb{R} : u \in \mathcal{A}_{L_c(W)}^\infty\} = \begin{cases} J_W^{\mathrm{rlx}}(u) & \text{for } u \in L^\infty(\Omega), \\ \infty & \text{otherwise,} \end{cases}$$

which is (6.4). Note that the last identity uses Lemma 5.1 and the simple fact that $u \notin \mathcal{A}_{L_c(W)}^\infty$ for any $c \in \mathbb{R}$ if $u \notin L^\infty(\Omega; \mathbb{R}^m)$.

Step 3: Upper bound. Let $u \in L^\infty(\Omega; \mathbb{R}^m)$. The construction of a recovery sequence $(u_j)_j \subset L^q(\Omega; \mathbb{R}^m)$ with the properties that $u_j \rightharpoonup u$ in $L^q(\Omega; \mathbb{R}^m)$ as $j \to \infty$ and

$$\limsup_{j \to \infty} I_W^{p_j}(u_j) \leq J_W^{\mathrm{rlx}}(u)$$

is straightforward. Indeed, the definition of the relaxed functional in (1.3) provides a sequence $(u_k)_k \subset L^\infty(\Omega; \mathbb{R}^m)$ such that $u_k \rightharpoonup^* u$ in $L^\infty(\Omega; \mathbb{R}^m)$ as $k \to \infty$ and

$$\lim_{k \to \infty} J_W(u_k) = J_W^{\mathrm{rlx}}(u).$$

Since $W(v_{u_k})$ is essentially bounded on $\Omega \times \Omega$ due to the growth assumption on W in (1.6), (2.6) and (5.1) imply

$$\lim_{j \to \infty} I_W^{p_j}(u_k) = \lim_{j \to \infty} \|W(v_{u_k})\|_{L^{p_j}(\Omega \times \Omega)} = \|W(v_{u_k})\|_{L^\infty(\Omega \times \Omega)} = J_W(u_k);$$

recall the notation in (4.11). Hence, in the limit $k \to \infty$,

$$\lim_{k \to \infty} \lim_{j \to \infty} I_W^{p_j}(u_k) = J_W^{\mathrm{rlx}}(u),$$

and we conclude by taking a diagonal sequence in the sense of Attouch, cf. [6, Corollary 1.16]. $\square$

Acknowledgments

Appendix. Basic technical tools

We collect here a few auxiliary results that are useful for our analysis of Cartesian convex hulls.

The following sequential compactness for sequences of compact sets with respect to the Hausdorff distance is a version of a classical statement (see e.g. [29, Section 3.9]) tailored to our needs.

Lemma A.1. *Let $(K_j)_j \subset \mathbb{R}^m \times \mathbb{R}^m$ be a sequence of compact sets with $K_{j+1} \subset K_j$ for all $j \in \mathbb{N}$ and $K := \bigcap_{j \in \mathbb{N}} K_j$. For any sequence of nonempty compact sets $(A_j)_j \subset \mathbb{R}^m$ such that $A_j \times A_j \subset K_j$ for all $j \in \mathbb{N}$, there exists a nonempty compact set $A \subset \mathbb{R}^m$ satisfying $A \times A \subset K$ and*

$$\liminf_{j \to \infty} d_H(A_j^{\mathrm{co}}, A^{\mathrm{co}}) = 0.$$

Proof. Since the projection onto the first vector variable of K_1, denoted by $\pi_1(K_1)$, is compact and $A_j \subset \pi_1(K_1)$ for all $j \in \mathbb{N}$, we exploit the compactness of the metric space of nonempty compact subsets of a compact set in $\mathbb{R}^m$ equipped with the Hausdorff distance to find a compact set $A \subset \pi_1(K_1)$ as the Hausdorff limit of a (non-relabeled) subsequence of $(A_j)_j$, that is,

$$d_H(A_j, A) \to 0 \quad \text{as } j \to \infty. \tag{A.1}$$

A well-known result [29, Theorem 3.9.4] on the relation between Hausdorff distances and convexifications states that any two nonempty and compact sets $C, D \subset \mathbb{R}^m$ satisfy

$$d_H(C^{\mathrm{co}}, D^{\mathrm{co}}) \leq d_H(C, D).$$

Applied to (A.1), this yields $d_H(A_j^{\mathrm{co}}, A^{\mathrm{co}}) \to 0$ as $j \to \infty$.

To show that $A \times A \subset K$, one observes that for any $(\xi, \zeta) \in A \times A$ and $j \in \mathbb{N}$,

$$\operatorname{dist}\big((\xi, \zeta), K\big) \leq \operatorname{dist}\big((\xi, \zeta), K_j\big) + d_H(K_j, K) \leq d_H(A \times A, A_j \times A_j) + d_H(K_j, K)$$
$$\leq d_H(A_j, A) + d_H(K_j, K).$$

Together with (A.1) and $\lim_{j \to \infty} d_H(K_j, K) = 0$, which follows from the fact that the limit in Hausdorff distance of a sequence of nested nonempty compact sets is exactly their intersection, we conclude that $(\xi, \zeta) \in K$. $\square$

The next lemma can be seen as the counterpart of [34, Lemma 4.9] for Cartesian convex hulls in place of separately convex ones.

Lemma A.2. *Let $(K_j)_j \subset \mathbb{R}^m \times \mathbb{R}^m$ be a sequence of nonempty, symmetric, diagonal, and compact sets with $K_{j+1} \subset K_j$ for all $j \in \mathbb{N}$. Then, $\bigcap_{j \in \mathbb{N}} K_j^{\times \mathrm{c}} = (\bigcap_{j \in \mathbb{N}} K_j)^{\times \mathrm{c}}$.*

Proof. Let $K := \bigcap_{j \in \mathbb{N}} K_j$. While $K^{\times c} = (\bigcap_{j \in \mathbb{N}} K_j)^{\times c} \subset \bigcap_{j \in \mathbb{N}} (K_j)^{\times c}$ is clear, the reverse inclusion follows based on the previous lemma and Lemma 3.4. Indeed, we show below that

$$\bigcap_{j \in \mathbb{N}} (K_j)_k^{\times c} \subset K_k^{\times c} \tag{A.2}$$

for any $k \in \mathbb{N}$. Then, Lemma 3.4 allows us to infer

$$\bigcap_{j \in \mathbb{N}} K_j^{\times c} = \bigcap_{j \in \mathbb{N}} \bigcup_{k \in \mathbb{N}} (K_j)_k^{\times c} = \bigcup_{k \in \mathbb{N}} \bigcap_{j \in \mathbb{N}} (K_j)_k^{\times c} \subset \bigcup_{k \in \mathbb{N}} K_k^{\times c} = K^{\times c},$$

which yields the statement.

We prove (A.2) via induction. Starting with $k = 1$, let $(\xi, \zeta) \in (K_j)_1^{\times c}$ for all $j \in \mathbb{N}$. Then there exists for each $j \in \mathbb{N}$ a set $B_j \subset \mathbb{R}^m$ with

$$B_j \times B_j \in \mathcal{P}_{K_j} = \mathcal{P}_{(K_j)_0^{\times c}} \quad \text{and} \quad \xi, \zeta \in B_j^{\text{co}}.$$

Notice that all these B_j are compact as maximal Cartesian squares of compact sets. According to Lemma A.1, there is then a nonempty compact set $B \subset \mathbb{R}^m$ such that $B \times B \subset K = \bigcap_{j \in \mathbb{N}} (K_j)_0^{\times c}$ and

$$\max\{\text{dist}(\xi, B^{\text{co}}), \text{dist}(\zeta, B^{\text{co}})\} \leq \liminf_{j \to \infty} d_H(B_j^{\text{co}}, B^{\text{co}}) = 0.$$

This shows $(\xi, \zeta) \in B^{\text{co}} \times B^{\text{co}} \subset K_1^{\times c}$. Repeating the same argument successively for $k \in \mathbb{N}$, taking into account Lemma 3.6, yields (A.2). $\square$

For an arbitrary sequence of sets $(E_j)_j \subset \mathbb{R}^m \times \mathbb{R}^m$ it holds that

$$\bigcap_{j \in \mathbb{N}} \mathcal{A}_{E_j} = \mathcal{A}_{\bigcap_{j \in \mathbb{N}} E_j},$$

see [34, Lemma 5.2]. We provide an asymptotic version of this identity when the sets are compact and nested.

Lemma A.3. *Let* $(K_j)_j \subset \mathbb{R}^m \times \mathbb{R}^m$ *be a sequence of nonempty, symmetric, diagonal, and compact sets with* $K_{j+1} \subset K_j$ *for all* $j \in \mathbb{N}$*, and let* $K := \bigcap_{j \in \mathbb{N}} K_j$*. Then,*

$$\bigcap_{j \in \mathbb{N}} \mathcal{A}_{K_j}^\infty = \mathcal{A}_K^\infty.$$

Proof. To address the nontrivial inclusion, take $u \in \bigcap_{j \in \mathbb{N}} \mathcal{A}_{K_j}^\infty$. In light of the representation formula (4.3), there exists a sequence $(B_j)_j \subset \mathbb{R}^m$ with $B_j \times B_j \in \mathcal{P}_{K_j}$ for all $j \in \mathbb{N}$ such that

$$u \in \bigcap_{j \in \mathbb{N}} L^\infty(\Omega; B_j^{\text{co}}) = L^\infty(\Omega; \bigcap_{j \in \mathbb{N}} B_j^{\text{co}}).$$

We can use Lemma A.1 to infer the existence of a nonempty compact set $B \subset \mathbb{R}^m$ with $B \times B \subset K$ and

$$\bigcap_{j \in \mathbb{N}} B_j^{\text{co}} \subset B^{\text{co}};$$

the details of this argument are similar to those in the proof of (A.2). Then, $u \in L^\infty(\Omega; B^{\text{co}}) \subset \mathcal{A}_K^\infty$, exploiting again (4.3). $\square$

References

[1] E. Acerbi, G. Buttazzo, F. Prinari, The class of functionals which can be represented by a supremum, J. Convex Anal. 9 (1) (2002) 225–236.

[2] E. Acerbi, N. Fusco, Semicontinuity problems in the calculus of variations, Arch. Ration. Mech. Anal. 86 (2) (1984) 125–145.

[3] N. Ansini, F. Prinari, Power-law approximation under differential constraints, SIAM J. Math. Anal. 46 (2) (2014) 1085–1115.

[4] H. Antil, Z.W. Di, R. Khatri, Bilevel optimization, deep learning and fractional laplacian regularization with applications in tomography, Inverse Problems 36 (6) (2020) 064001.

[5] G. Aronsson, Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, I-III, Ark. Mat. 6 (1965) 33–53.

[6] H. Attouch, Variational Convergence for Functions and Operators, in: Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[7] R.J. Aumann, S. Hart, Bi-convexity and bi-martingales, Israel J. Math. 54 (2) (1986) 159–180.

[8] E.N. Barron, R.R. Jensen, C.Y. Wang, The Euler equation and absolute minimizers of L^∞ functionals, Arch. Ration. Mech. Anal. 157 (4) (2001) 255–283.

[9] E.N. Barron, R.R. Jensen, C.Y. Wang, Lower semicontinuity of L^∞ functionals, Ann. Inst. H. Poincaré Anal. Non Linéaire 18 (4) (2001) 495–517.

[10] E.N. Barron, W. Liu, Calculus of variations in L^∞, Appl. Math. Optim. 35 (3) (1997) 237–263.

[11] J.C. Bellido, J. Cueto, C. Mora-Corral, Bond-based peridynamics does not converge to hyperelasticity as the horizon goes to zero, J. Elast. 141 (2) (2020) 273–289.

[12] J.C. Bellido, J. Cueto, C. Mora-Corral, Fractional piola identity and polyconvexity in fractional spaces, Ann. Inst. H. Poincaré Anal. Non Linéaire 37 (4) (2020) 955–981.

[13] J.C. Bellido, C. Mora-Corral, Lower semicontinuity and relaxation via Young measures for nonlocal variational problems and applications to peridynamics, SIAM J. Math. Anal. 50 (1) (2018) 779–809.

[14] J. Bevan, P. Pedregal, A necessary and sufficient condition for the weak lower semicontinuity of one-dimensional non-local variational integrals, Proc. Roy. Soc. Edinburgh Sect. A 136 (4) (2006) 701–708.

[15] H. Brezis, H.-M. Nguyen, Non-local functionals related to the total variation and connections with image processing, Ann. PDE 4 (1) (2018) Paper No. 9, 77.

[16] A. Briani, F. Prinari, A. Garroni, Homogenization of L^∞ functionals, Math. Models Methods Appl. Sci. 14 (12) (2004) 1761–1784.

[17] G. Buttazzo, G. Dal Maso, On nemyckii operators and integral representation of local functionals, Rend. Mat. (7) 3 (3) (1983) 491–509.

[18] P. Cardaliaguet, F. Prinari, Supremal representation of L^∞ functionals, Appl. Math. Optim. 52 (2) (2005) 129–141.

[19] T. Champion, L. De Pascale, F. Prinari, Γ-convergence and absolute minimizers for supremal functionals, ESAIM Control Optim. Calc. Var. 10 (1) (2004) 14–27.

[20] B. Dacorogna, Direct Methods in the Calculus of Variations, second ed., in: Applied Mathematical Sciences, vol. 78, Springer, New York, 2008.

[21] G. Dal Maso, An Introduction to Γ-Convergence, in: Progress in Nonlinear Differential Equations and their Applications, vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1993.

[22] G. Dal Maso, I. Fonseca, G. Leoni, Asymptotic analysis of second order nonlocal Cahn-Hilliard-type functionals, Trans. Amer. Math. Soc. 370 (4) (2018) 2785–2823.

[23] M. D'Elia, J.C.D. los Reyes, A.M. Trujillo, Bilevel parameter optimization for learning nonlocal image denoising models, 2020.

[24] G. Dolzmann, Variational Methods for Crystalline Microstructure—Analysis and Computation, in: Lecture Notes in Mathematics, vol. 1803, Springer-Verlag, Berlin, 2003.

[25] M. Eleuteri, F. Prinari, Γ-convergence for power-law functionals with variable exponents, Nonlinear Anal. RWA 58 (2021) 103221, 21.

[26] I. Fonseca, G. Leoni, Modern Methods in the Calculus of Variations: L^P Spaces, in: Springer Monographs in Mathematics, Springer, New York, 2007.

[27] I. Fonseca, S. Müller, $\mathcal{A}$-quasiconvexity, lower semicontinuity, and Young measures, SIAM J. Math. Anal. 30 (6) (1999) 1355–1390.

[28] A. Garroni, V. Nesi, M. Ponsiglione, Dielectric breakdown: optimal bounds, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. 457 (2014) (2001) 2317–2335.

[29] V.V. Goncharov, Anàlise Multìvoca e Nao Suave, Universidade de Èvora, 2018.

[30] G. Holler, K. Kunisch, Learning nonlocal regularization operators, 2020.

[31] N. Katzourakis, An Introduction to Viscosity Solutions for Fully Nonlinear PDE with Applications to Calculus of Variations in L^∞, in: SpringerBriefs in Mathematics, Springer, Cham, 2015.

[32] J. Kolář, Non-compact lamination convex hulls, Ann. Inst. H. Poincaré Anal. Non Linéaire 20 (3) (2003) 391–403.

[33] C. Kreisbeck, H. Schönberger, Quasiconvexity in the fractional calculus of variations: characterization of lower semicontinuity and relaxation, Nonlinear Anal. 215 (2022) Paper No. 112625, 26.

[34] C. Kreisbeck, E. Zappale, Lower semicontinuity and relaxation of nonlocal L^∞-functionals, Calc. Var. Partial Differential Equations 59 (4) (2020) Paper No. 138, 36.

[35] C. Kreisbeck, E. Zappale, Loss of double-integral character during relaxation, SIAM J. Math. Anal. 53 (1) (2021) 351–385.

[36] T. Mengesha, Q. Du, On the variational limit of a class of nonlocal functionals related to peridynamics, Nonlinearity 28 (11) (2015) 3999–4035.

[37] C. Mora-Corral, A. Tellini, Relaxation of a scalar nonlocal variational problem with a double-well potential, Calc. Var. Partial Differential Equations 59 (2) (2020) Paper No. 67, 30.

[38] C.B. Morrey Jr., Multiple Integrals in the Calculus of Variations, in: Die Grundlehren der Mathematischen Wissenschaften, Band 130, Springer-Verlag New York, Inc., New York, 1966.

[39] J. Muñoz, Characterisation of the weak lower semicontinuity for a type of nonlocal integral functional: the n-dimensional scalar case, J. Math. Anal. Appl. 360 (2) (2009) 495–502.

[40] P. Pedregal, Parametrized Measures and Variational Principles, in: Progress in Nonlinear Differential Equations and their Applications, vol. 30, Birkhäuser Verlag, Basel, 1997.

[41] P. Pedregal, Nonlocal variational principles, Nonlinear Anal. 29 (12) (1997) 1379–1392.

[42] P. Pedregal, Weak lower semicontinuity and relaxation for a class of non-local functionals, Rev. Mat. Complut. 29 (3) (2016) 485–495.

[43] P. Pedregal, On non-locality in the calculus of variations, SeMA J. 78 (4) (2021) 435–456.

[44] F. Prinari, Relaxation and Γ-convergence of supremal functionals, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) 9 (1) (2006) 101–132.

[45] F. Prinari, Semicontinuity and relaxation of L^∞-functionals, Adv. Calc. Var. 2 (1) (2009) 43–71.

[46] F. Prinari, On the lower semicontinuity and approximation of L^∞-functionals, NoDEA Nonlinear Differential Equations Appl. 22 (6) (2015) 1591–1605.

[47] F. Prinari, E. Zappale, A relaxation result in the vectorial setting and power law approximation for supremal functionals, J. Optim. Theory Appl. 186 (2) (2020) 412–452.

[48] F. Rindler, Calculus of Variations, in: Universitext, Springer, Cham, 2018.

[49] O. Savin, E. Valdinoci, Γ-convergence for nonlocal phase transitions, Ann. Inst. H. Poincaré Anal. Non Linéaire 29 (4) (2012) 479–500.

[50] T.-T. Shieh, D.E. Spector, On a new class of fractional partial differential equations, Adv. Calc. Var. 8 (4) (2015) 321–336.