PROBABILITY INEQUALITIES FOR STRONGLY LEFT-INVARIANT METRIC SEMIGROUPS/MONOIDS, INCLUDING ALL LIE GROUPS

APOORVA KHARE

To the memory of K.R. Parthasarathy, with admiration

Abstract. Recently, a general version of the Hoffmann-Jørgensen inequality was shown jointly with Rajaratnam [*Ann. Probab.* 2017], which (a) improved the result even for real-valued variables, but also (b) simultaneously unified and extended several versions in the Banach space literature, including that by Hitczenko–Montgomery-Smith [*Ann. Probab.* 2001], as well as special cases and variants of results by Johnson–Schechtman [*Ann. Probab.* 1989] and Klass–Nowicki [*Ann. Probab.* 2000], in addition to the original versions by Kahane and Hoffmann-Jørgensen. Moreover, our result with Rajaratnam was in a primitive framework: over all semigroups with a bi-invariant metric; this includes Banach spaces as well as compact and abelian Lie groups.

In this note we show the result even more generally: over every semigroup $\mathscr{G}$ with a strongly left- (or right-)invariant metric. We also prove some applications of this inequality over such $\mathscr{G}$, extending Banach space-valued versions by Hitczenko and Montgomery-Smith [*Ann. Probab.* 2001] and by Hoffmann-Jørgensen [*Studia Math.* 1974]. Furthermore, we show several other stochastic inequalities – by Ottaviani–Skorohod, Mogul'skii, and Lévy–Ottaviani – as well as Lévy's equivalence, again over $\mathscr{G}$ as above. This setting of generality for $\mathscr{G}$ subsumes not only semigroups with bi-invariant metric (thus extending the previously shown results), but it also means that these results now hold over all Lie groups (equipped with a left-invariant Riemannian metric).

We also explain why this primitive setting of strongly left/right-invariant metric semigroups $\mathscr{G}$ is equivalent to that of left/right-invariant metric monoids $\mathscr{G}_\circ$: each such $\mathscr{G}$ embeds in some $\mathscr{G}_\circ$.

1. Introduction: strongly left/right-invariant metric semigroups

In this work our goal is to extend various results in the probability literature to very primitive settings – e.g. from real- or Banach space-valued random variables, to ones taking values in an arbitrary Lie group, and even more general classes of variables. This continues a series of recent joint works [17, 18, 19] by the author, which were inspired in part by the seminal treatise [25] of Parthasarathy that studied (probability) measures on very primitive structures: (separable) metric spaces, metric/locally compact abelian groups, and so on.

In this note, our goal is to extend stochastic inequalities, tail estimates, and convergence phenomena that were known to hold for random variables over Banach spaces, or – very recently – semigroups with bi-invariant metrics, to semigroups with strongly left-invariant (or right-invariant) metrics. We begin by defining these latter notions.

Definition 1.1.

(1) As defined in [17], a *bi-invariant metric semigroup* consists of a semigroup $(\mathscr{G}, \cdot)$ with a bi-invariant metric $d_{\mathscr{G}}$ – i.e.,

$$d_{\mathscr{G}}(ca, cb) = d_{\mathscr{G}}(a, b) = d_{\mathscr{G}}(ac, bc), \qquad \forall a, b, c \in \mathscr{G}. \tag{1.1}$$

Date: July 3, 2024.

1991 *Mathematics Subject Classification.* 60E15 (primary); 60B10 (secondary).

Key words and phrases. Metric semigroup, strongly left-invariant metric semigroup, left-invariant metric monoid, Hoffmann-Jørgensen inequality, Ottaviani–Skorohod inequality, Mogul'skii inequality, Lévy–Ottaviani inequality, Lévy equivalence, decreasing rearrangement, universal constant.

(In our previous joint works [17, 18, 19], we refer to such a $\mathscr{G}$ as merely a metric semigroup; however, the bi-invariance will be explicitly pointed out in this work, to distinguish from the following notions.)

(2) If only the first (respectively, second) equality in (1.1) holds for all $a, b, c \in \mathscr{G}$, then we say that $\mathscr{G}$ is a *left-* (respectively, *right-*)*invariant metric semigroup*.

(3) Similarly, one defines a *left/right/bi-invariant metric monoid/group*.

(4) Finally, a left-invariant metric semigroup $\mathscr{G}$ is *strongly left-invariant* if $d_{\mathscr{G}}(a, ab) = d_{\mathscr{G}}(b, b^2)$ for all $a, b \in \mathscr{G}$. One similarly defines a strongly right-invariant metric semigroup.

Remark 1.2. In other words, $d_{\mathscr{G}}$ is strongly left-invariant if and only if for all $b \in \mathscr{G}$, $d_{\mathscr{G}}(a, ab)$ is independent of $a \in \mathscr{G}$ (so one can set $a = b$ in $\mathscr{G}$). Thus every left-invariant metric monoid or group $(\mathscr{G}, \cdot, e, d_{\mathscr{G}})$ is a strongly left-invariant semigroup, since $d_{\mathscr{G}}(a, ab) = d_{\mathscr{G}}(e, b)$. (Ditto for right-invariant $d_{\mathscr{G}}$.) Similarly, it was shown in [17] (see (2.6) below) that every bi-invariant metric semigroup is strongly left- and right-invariant. Thus, adding an identity or right-invariance automatically upgrades left-invariance to its "strong" version, and so the present setting subsumes both of these related settings above/in previous works.

Remark 1.3. For $\mathscr{G}$ a group, we will not insist that $d_{\mathscr{G}}(a, b) = d_{\mathscr{G}}(a^{-1}, b^{-1})$ for all $a, b \in \mathscr{G}$. Note, if this holds then $d_{\mathscr{G}}$ is left/right-invariant if and only if it is bi-invariant, since e.g. one has:

$$d_{\mathscr{G}}(ac, bc) = d_{\mathscr{G}}((ac)^{-1}, (bc)^{-1}) = d_{\mathscr{G}}(c^{-1}a^{-1}, c^{-1}b^{-1}) = d_{\mathscr{G}}(a^{-1}, b^{-1}) = d_{\mathscr{G}}(a, b), \qquad \forall a, b, c \in \mathscr{G}.$$

The goal of this work is to extend results from random variables taking values in Banach spaces (as is traditional by now) or even in bi-invariant metric semigroups (as was done in recent joint works), to strongly left/right-invariant metric semigroup-valued variables. The motivation to extend results from Banach spaces to more primitive frameworks is both classical and modern. Following its axiomatization and systematic development, one of the cornerstones of twentieth century probability theory has been to extend results for real-valued random variables to $\mathbb{B}$-valued random variables, for $\mathbb{B}$ a (separable) Banach space – see e.g. the classic treatise [21]. Now a natural theoretical question is to explore settings beyond Banach spaces. In fact such questions have been widely studied in the past few decades. We mention the classic monographs by Parthasarathy [25] and Grenander [7]; as well as (among many others) the Diaconis–Shahshahani work on random permutations [4] and the recent theory of (dense as well as sparse) graph limits – which has already been crystallized by Lovász in book form [22].[1] This activity continues to thrive; e.g. outside the (by now) traditional Banach space setting, we list a few of the many works involving random variables taking values in (possibly non-compact and non-abelian) Lie groups, in random matrix theory [2, 8, 14] and its connections to ergodic theory and geometry (see e.g. [5, 10, 26]).

Thus our goal in this note is to extend several results in the vast literature on Banach space-valued random variables – or more generally, results with variables valued in bi-invariant metric semigroups (these further include discrete, abelian, compact, or amenable groups, see the introduction to [17] for more examples and [17, 18, 19] for the results) – to an even more primitive setting: strongly left/right-invariant metric semigroups. In particular, since these include left-invariant metric groups, it follows that several results which were previously known only over compact or abelian Lie groups, now extend to every Lie group (including non-compact, non-abelian ones).

Remark 1.4. This work attempts to provide the *most primitive setting* in which the results stated below can be proved. To this end:

(1) By Remark 1.2, random variables taking values in strongly left/right-invariant metric semigroups $\mathscr{G}_{strong}$ subsume those taking values in either left/right-invariant metric monoids $\mathscr{G}_\circ$ or in bi-invariant metric semigroups $\mathscr{G}$. So working with such semigroups $\mathscr{G}_{strong}$ is at least as general.

[1] While every metric space isometrically embeds into a Banach space by the Kuratowski embedding theorem, the study of graphons with the cut-norm does not typically proceed using this embedding.

(2) While the results below can be stated over any semigroup with a metric, it is not clear (to the author) how to prove them in a more primitive setting than strong left/right-invariance, since it is indispensable in the proofs of most/all of the results below. One way to avoid this technical hurdle (the word "strong") could be if – as in the bi-invariant case – that every left/right-invariant metric semigroup embeds isometrically and homomorphically inside a left/right-invariant metric monoid $\mathscr{G}_\circ$, because then $d_{\mathscr{G}}(a, ab) = d_{\mathscr{G}_\circ}(e, b) = d_{\mathscr{G}}(b, b^2)$. As we show in Proposition 5.2, this does not always happen.

Thus, we suspect that the strong left- or right-invariance of the metric is perhaps the *minimum* amount of structure required in order to be able to show the Hoffmann-Jørgensen, Ottaviani–Skorohod, Mogul'skii, and Lévy–Ottaviani inequalities and their applications shown in this work.

We end here with the punchline of the "non-probabilistic part" of the paper: *every strongly left/right-invariant metric semigroup $\mathscr{G}_{strong}$ embeds isometrically and homomorphically into a left/right-invariant metric monoid $(\mathscr{G}_\circ, e)$.* (In fact, we show that $\mathscr{G}_{strong} \supseteq \mathscr{G}_\circ \setminus \{e\}$, so that the smallest $\mathscr{G}_\circ \supseteq \mathscr{G}_{strong}$ is unique; see Theorem 5.1.) Thus, our suspected "most primitive setting" for proving stochastic inequalities, of working with $\mathscr{G}_{strong}$-valued random variables, is actually *equivalent* to working with $\mathscr{G}_\circ$-valued random variables. This equivalence parallels Proposition 2.10 (drawn from recent joint work [18]), which gives a similar equivalence for bi-invariant metric semigroups vs. monoids.

2. The Hoffmann-Jørgensen inequality for strongly left-invariant metric semigroups

The first inequality that we extend to strongly left/right-invariant metric semigroups is the Hoffmann-Jørgensen inequality, which is used to bound sums of independent random variables. We refer the reader to e.g. [9] for a detailed history of the inequality.

2.1. The inequality over strongly left/right-invariant metric, following previous variants.

Here we will present a few versions of the above inequality from the literature, ending with a general, unifying variant that holds over all bi-invariant metric semigroups, before extending it (and hence the preceding variants) to the more general setting of strongly left/right-invariant metric semigroups. We start with a version found in the monograph by Ledoux–Talagrand, where the authors attribute the result to Hoffmann-Jørgensen [11] (see also Kahane [16]).

Theorem 2.1 ([21, Proposition 6.7])**.** *Suppose $\mathbb{B}$ is a separable Banach space, and $(\Omega, \mathscr{A}, \mu)$ is a probability space with $X_1, \ldots, X_n \in L^0(\Omega, \mathbb{B})$ independent random variables. For $1 \leqslant j \leqslant n$, define $S_j := X_1 + \cdots + X_j$ and $U_n := \max_{1 \leqslant j \leqslant n} \|S_j\|$. Then*

$$\mathbb{P}_\mu\left(U_n > 3t + s\right) \leqslant \mathbb{P}_\mu\left(U_n > t\right)^2 + \mathbb{P}_\mu\left(\max_{1 \leqslant j \leqslant n} \|X_j\| > s\right), \qquad \forall s, t \in (0, \infty).$$

Theorem 2.1 has led to several strengthenings and variants, including by Johnson–Schechtman [15], Klass–Nowicki [20], and Hitczenko and Montgomery-Smith [9]. This last variant says:

Theorem 2.2 ([9, Theorem 1])**.** *(Notation as in Theorem 2.1.) For all $K \in \mathbb{Z}_{>0}$ and $s, t \in (0, \infty)$,*

$$\mathbb{P}_\mu\left(U_n > 2Kt + (K-1)s\right) \leqslant \frac{1}{K!}\left(\frac{\mathbb{P}_\mu\left(U_n > t\right)}{\mathbb{P}_\mu\left(U_n \leqslant t\right)}\right)^K + \mathbb{P}_\mu\left(\max_{1 \leqslant j \leqslant n} \|X_j\| > s\right).$$

Notice from their statements that neither of Theorems 2.1 and 2.2 immediately follow from the other. They were both simultaneously extended in recent work [17], by the next variant that we state. To do so, we require some notation.

Definition 2.3. Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable semigroup with a left-invariant metric, with Borel σ-algebra $\mathscr{B}_{\mathscr{G}}$. Given integers $1 \leqslant j \leqslant n$ and random variables $X_1, \ldots, X_n : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_{\mathscr{G}})$, define

$$S_j(\omega) := X_1(\omega) \cdots X_j(\omega), \qquad M_j(\omega) := \max_{1 \leqslant i \leqslant j} d_{\mathscr{G}}(z_0, z_0 X_i(\omega)), \tag{2.1}$$

where $z_0 \in \mathscr{G}$ is arbitrary.

Remark 2.4. Clearly, M_j is independent of $z_0 \in \mathscr{G}$ in any strongly left/right-invariant metric semigroup, hence – as observed above – in any left/right-invariant metric monoid. We explain presently why this implies the same fact over every bi-invariant metric semigroup as in [17].

With this notation at hand, we state the next – and rather general – variant of the Hoffmann-Jørgensen inequality:

Theorem 2.5 ([17, Theorem A])**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable bi-invariant metric semigroup, $z_0, z_1 \in \mathscr{G}$ are fixed, and $X_1, \ldots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. Also fix integers $0 < k, n_1, \ldots, n_k \in \mathbb{Z}$ and nonnegative scalars $t_1, \ldots, t_k, s \in [0, \infty)$, and define*

$$U_n := \max_{1 \leqslant j \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_j), \quad I_0 := \{1 \leqslant i \leqslant k : \mathbb{P}_\mu (U_n \leqslant t_i)^{n_i - \delta_{i1}} \leqslant \frac{1}{n_i!}\}, \tag{2.2}$$

where δ_{i1} denotes the Kronecker delta. Now if $\sum_{i=1}^k n_i \leqslant n+1$, then

$$\begin{aligned} &\mathbb{P}_\mu \left(U_n > (2n_1 - 1)t_1 + 2\sum_{i=2}^k n_i t_i + \left(\sum_{i=1}^k n_i - 1 \right) s \right) \\ &\leqslant \mathbb{P}_\mu (U_n \leqslant t_1)^{\mathbf{1}_{1 \notin I_0}} \prod_{i \in I_0} \mathbb{P}_\mu (U_n > t_i)^{n_i} \prod_{i \notin I_0} \frac{1}{n_i!} \left(\frac{\mathbb{P}_\mu (U_n > t_i)}{\mathbb{P}_\mu (U_n \leqslant t_i)} \right)^{n_i} \\ &\quad + \mathbb{P}_\mu (M_n > s) . \end{aligned} \tag{2.3}$$

More generally, define

$$\begin{aligned} &K := \sum_{i=1}^k n_i, \qquad Y_j := d_{\mathscr{G}}(z_0, z_0 X_j), \\ &Y_{(1)} := \min(Y_1, \ldots, Y_n), \quad \ldots, \quad Y_{(n)} := \max(Y_1, \ldots, Y_n), \end{aligned}$$

so that $Y_{(j)}$ are the order statistics of the Y_j. Then the above inequality can be strengthened by replacing $\mathbb{P}_\mu (M_n > s)$ by

$$\mathbb{P}_\mu \left(\sum_{j=n-K+2}^n Y_{(j)} > (K-1)s \right) .$$

It is this result whose setting we weaken, to obtain the main result of this section:

Theorem 2.6. *Theorem 2.5 holds more generally, over every strongly left/right-invariant metric semigroup (equivalently by Theorem 5.1, over every left/right-invariant metric monoid).*

As explained in [17], Theorem 2.6 specializes to Theorem 2.1 by setting

$$\mathscr{G} = \mathbb{B}, \quad z_0 = z_1 = 0, \quad k = 2, \quad n_1 = n_2 = 1, \quad t_1 = t_2 = t,$$

so that $I_0 = \{1, 2\}$. It also specializes to Theorem 2.2, by setting $\mathscr{G} = \mathbb{B},\ z_0 = z_1 = 0,\ k = 1,\ n_1 = K,\ t_1 = t$. But moreover, Theorem 2.6 specializes to Theorem 2.5 itself (which implied the other two results above) via Proposition 2.10 below: every bi-invariant metric semigroup embeds into a bi-invariant metric monoid, hence is a strongly left/right-invariant metric semigroup.

2.2. **Additional previous variants.** The proof of Theorem 2.6 is deferred to the next subsection, together with the fact that it implies Theorem 2.5. For now, we elaborate in some detail the connection of Theorems 2.5 and 2.6 to (special cases of) results by Johnson–Schechtman [15] and Klass–Nowicki [20]. First, Johnson–Schechtman showed by iterating the "original" Hoffmann-Jørgensen inequality:

Theorem 2.7 ([15, Lemma 6])**.** *If $X_1, \dots, X_n$ are independent non-negative real-valued random variables, and $z_0 = z_1 = 0$ (so that $S_n = \|S_n\| = U_n$), then*

$$\mathbb{P}_\mu\left(U_n > (2k-1)t\right) \leqslant \mathbb{P}_\mu\left(M_n > t\right) + \mathbb{P}_\mu\left(U_n > t\right)^k, \qquad \forall t \in (0,\infty),\ k \in \mathbb{Z}_{>0}.$$

Theorem 2.5 (and hence Theorem 2.6) implies a weaker form of this result, in which $\mathbb{P}_\mu\left(U_n > t\right)$ is replaced by $\mathbb{P}_\mu\left(U_n > t/2\right)$. To see this, set

$$\mathscr{G} = \mathbb{R}, \quad z_0 = z_1 = 0, \quad n_1 = \cdots = n_k = 1, \quad t_1 = s = t,\ t_2 = \cdots = t_k = t/2,$$

so that $I_0 = \{1, \dots, k\}$. Also note that Theorem 2.7 holds for non-negative variables, so $\mathscr{G}$ is necessarily $\mathbb{R}$, while Theorem 2.6 holds for all $\mathscr{G}$ – including possibly negative real-valued random variables – and its assertion in this generality is slightly weaker.

Second, Hitczenko and Montgomery-Smith term their result (Theorem 2.2) the *Klass–Nowicki inequality*. This is because Klass–Nowicki had previously shown a slightly different – but related – result to Theorem 2.2:

Theorem 2.8 ([20, Theorem 1.1])**.** *Suppose $X_1, \dots, X_n$ are independent Banach space-valued random variables, and we define S_n, U_n as above with $z_0 = z_1 = 0$. Also set $\lambda := \mathbb{P}_\mu\left(U_n \geqslant 1\right)$ and suppose $\lambda < 1$. If the X_j are non-negative, then*

$$\mathbb{P}_\mu\left(\|S_n\| \geqslant k + Y_{(n)} + \cdots + Y_{(n-k+2)}\right) \leqslant \frac{1}{k!}[n(1 - \sqrt[n]{1-\lambda})]^k, \qquad \forall k \geqslant 1, \tag{2.4}$$

whereas if the X_j are symmetric, then

$$\mathbb{P}_\mu\left(U_n \geqslant k + Y_{(n)} + \cdots + Y_{(n-k+2)}\right) \leqslant \frac{2^{k-1}}{k!}[n(1 - \sqrt[n]{1-\lambda})]^k, \qquad \forall k \geqslant 1. \tag{2.5}$$

Note that the left-hand expression in (2.4) is bounded above by the one in (2.5). These results relate to Theorems 2.5 and 2.6 with $t = s = 1$ – even when the X_j need not be non-negative or symmetric – as follows:

$$\begin{aligned}
&\mathbb{P}_\mu\left(U_n > 2kt + (k-1)s\right) \\
&\leqslant \mathbb{P}_\mu\left(U_n > kt + (k-1)s\right) \\
&\leqslant \mathbb{P}_\mu\left(U_n > k + Y_{(n)} + \cdots + Y_{(n-k+2)}\right) + \mathbb{P}_\mu\left(Y_{(n)} + \cdots + Y_{(n-k+2)} > (k-1)\right) \\
&\leqslant \frac{2^{k-1}}{k!}[n(1 - \sqrt[n]{1-\lambda})]^k + \mathbb{P}_\mu\left(Y_{(n)} + \cdots + Y_{(n-k+2)} > (k-1)\right),
\end{aligned}$$

where the final inequality uses (2.5). Now if the upper bound in the final inequality was missing the factor of 2^{k-1}, then we **claim** that this bound is at most $\frac{1}{k!}(\lambda/(1-\lambda))^k$, which is the precise factor in Theorem 2.2 and hence in a suitable specialization of Theorem 2.5. Thus, Theorem 2.8 would imply a very similar result to Theorem 2.2 (strengthened to replace $\mathbb{P}_\mu\left(M_n > 1\right)$ by $\mathbb{P}_\mu\left(Y_{(n)} + \cdots + Y_{(n-k+2)} > (k-1)\right)$, as in [9] as well as in Theorem 2.5).

Remark 2.9. Here we quickly explain why the result obtained in the preceding paragraph would be similar to Theorem 2.2 but not exactly comparable. The preceding claim – that the bound in (2.4) is lower than the bound in Theorem 2.2 – is easily checked:

$$\frac{1}{k!}[n(1 - \sqrt[n]{1-\lambda})]^k \leqslant \frac{1}{k!}\left(\frac{\lambda}{1-\lambda}\right)^k, \qquad \forall k \geqslant 1.$$

This is because the inequality turns out to be equivalent to requiring $1-\alpha\lambda \leqslant (1-\lambda)^\alpha$ for $\lambda \in [0,1)$ and $\alpha = 1 + \frac{1}{n} \in [1,2]$, and this latter inequality holds by the binomial series formula. However, the same inequality does not hold when working with the bound in (2.5) instead:

$$\frac{2^{k-1}}{k!}[n(1-\sqrt[n]{1-\lambda})]^k \not\leqslant \frac{1}{k!}\left(\frac{\lambda}{1-\lambda}\right)^k, \qquad \forall k > 1.$$

Indeed, this non-inequality $\not\leqslant$ is equivalent to $>$, i.e., to the assertion that

$$2^{1-1/k} n(1-\lambda)(1-\sqrt[n]{1-\lambda}) > \lambda,$$

which does hold at very small values of $\lambda \in (0,1)$, since the left-hand side is a power series in λ with constant term zero and linear term $2^{1-1/k} > 1$ (for $k > 1$).

2.3. **The proofs.** Before showing Theorem 2.6, let us prove that it implies Theorem 2.5: this is because every metric semigroup embeds into a bi-invariant metric monoid, by the following result.

Proposition 2.10 ([19, §2.1])**.** *Every bi-invariant metric semigroup $\mathscr{G}$ is contained in a metric monoid. More precisely, $\mathscr{G}$ contains at most one idempotent, which is automatically a two-sided identity. Thus if $\mathscr{G}$ contains exactly one idempotent then it is a metric monoid. Otherwise, $\mathscr{G}$ is the set of non-identity elements in a metric monoid $\mathscr{G}_\circ$ (with identity e; thus the smallest $\mathscr{G}_\circ \supseteq \mathscr{G}$ is unique), and $\mathscr{G}$ maps isometrically into $\mathscr{G}_\circ$ with: $d_{\mathscr{G}_\circ}(e,g) := d_{\mathscr{G}}(g,g^2)$ for all $g \in \mathscr{G}$.*

The proof is not hard (but omitted here; see [19]). Note that the final statement about the extended metric being bi-invariant uses the following calculation in any metric semigroup:

$$d_{\mathscr{G}}(a,ba) = d_{\mathscr{G}}(ba,b^2a) = d_{\mathscr{G}}(b,b^2) = d_{\mathscr{G}}(ab,ab^2) = d_{\mathscr{G}}(a,ab), \qquad \forall a,b \in \mathscr{G}. \tag{2.6}$$

Of course, the equalities in (2.6) need not all hold in every left/right-invariant metric monoid (such as in non-compact non-abelian Lie groups G, which possess left-invariant but not necessarily bi-invariant Riemannian metrics[2]).

The other proof is that of our main result in this section; this is now provided and includes fixing a small typo in [17] itself.

Proof of Theorem 2.6. The idea is to closely follow the proof of Theorem 2.5, making the necessary adjustments along the way because now $d_{\mathscr{G}}$ is only strongly left-invariant. The first change is that the assertion that $Y_j := d_{\mathscr{G}}(z_0, z_0X_j)$, $M_j(\omega) := \max_{1\leqslant i\leqslant j} Y_j$ do not depend on z_0, which was shown in [17] using (2.6) via bi-invariance, now follows instead from the strong left-invariance of $d_{\mathscr{G}}$. The next change involves avoiding the use of (2.6) in Step 2 of the proof, in the long calculation on pp. 4106 of [17], where we computed:

$$\begin{aligned} d_{\mathscr{G}}(z_0S_{m-1}(\omega), z_0S_m(\omega)) = d_{\mathscr{G}}(z_0S_{m-1}(\omega), z_0S_{m-1}(\omega)\cdot X_m(\omega)) &= d_{\mathscr{G}}(z_0, z_0X_m(\omega)) \\ &= Y_m(\omega). \end{aligned}$$

Instead, this calculation now holds using merely the strong left-invariance of $d_{\mathscr{G}}$.

A small remark here concerns an argument in [9, Proof of Theorem 1] over Banach spaces, that was black-boxed in [17]:

$$\sup_{j,k\leqslant n} \|S_k - S_j\| \leqslant 2\sup_j \|S_j\| = 2U_n.$$

This was used in [17] for bi-invariant metric semigroups; as we now note, one requires only the (not even strong) left-invariance of $d_{\mathscr{G}}$, and not an identity in $\mathscr{G}$, since by the triangle inequality,

$$\sup_{j,k\leqslant n} d_{\mathscr{G}}(S_k,S_j) = \max_{1\leqslant j,k\leqslant n} d_{\mathscr{G}}(z_0S_k, z_0S_j) \leqslant 2\max_{1\leqslant j\leqslant n} d_{\mathscr{G}}(z_1, z_0S_j) = 2U_n.$$

[2]For instance, by [24], the only connected Lie groups that admit a bi-invariant Riemannian metric are of the form $K \times \mathbb{R}^n$, for K a compact group and $n \geqslant 0$ an integer.

The final point here is a modification that in fact applies to the *original* proof itself in [17] – we fix a small typo there. Namely, in [17, Equation (9)], the definition of $p_{\beta,t}$ should be modified to use $0 < j$ instead of $0 \leqslant j$. Thus we need to define and work with

$$p_{\beta,t} := \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_1, z_0 S_\beta) > t \geqslant d_{\mathscr{G}}(z_1, z_0 S_j) \ \forall 0 < j < \beta \right), \qquad 0 < \beta \leqslant n. \tag{2.7}$$

□

Remark 2.11. We leave it to the interested reader to work out the case of strongly right-invariant metric semigroups, for the results in this section and the next. For this, when working only over semigroups, the results and proofs involve the modified quantities

$$S_j(\omega) := X_j(\omega) X_{j-1}(\omega) \cdots X_1(\omega), \qquad U_n := \max_{1 \leqslant j \leqslant n} d_{\mathscr{G}}(z_1, S_j z_0), \qquad Y_j := d_{\mathscr{G}}(z_0, X_j z_0).$$

However, if $\mathscr{G}$ is in fact a group then one does not need to prove the "strongly right-invariant" analogues (of all results in this paper) separately, since it follows from the left-invariant results in Theorem 2.6. This is because of the bijection between the sets of left- and right-invariant metrics on a group G: $d^L_G(a,b) \iff d^R_G(a^{-1}, b^{-1})$. The same holds if G is abelian, since any left/right-invariant metric d_G is automatically bi-invariant, and the results in this work reduce to those in previous works [17, 18, 19].

3. Decreasing rearrangements and applications

Here we write down a couple of applications of the Hoffmann-Jørgensen inequality. The first shows that controlling the behavior of the independent variables X_j is the same as controlling S_n or U_n – but now over all strongly left/right-invariant semigroups (equivalently by Theorem 5.1, over left/right-invariant metric monoids):

Theorem 3.1. *Suppose $A \subseteq \mathbb{Z}_{>0}$ is either $\mathbb{Z}_{>0}$ or $\{1, \dots, N\}$ for some $N \in \mathbb{Z}_{>0}$. Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable strongly left-invariant metric semigroup, $z_0, z_1 \in \mathscr{G}$, and the variables $X_n \in L^0(\Omega, \mathscr{G})$ are independent for all $n \in A$. If $\sup_{n \in A} d_{\mathscr{G}}(z_1, z_0 S_n) < \infty$ almost surely, then for all $p \in (0, \infty)$,*

$$\mathbb{E}_\mu \left[\sup_{n \in A} d_{\mathscr{G}}(z_0, z_0 X_n)^p \right] < \infty \quad \iff \quad \mathbb{E}_\mu \left[\sup_{n \in A} d_{\mathscr{G}}(z_1, z_0 S_n)^p \right] < \infty.$$

This extends [18, Theorem A] and originally [11, Theorem 3.1] to the primitive setting of strongly left-invariant metric semigroups. The proof of this and the next few results use the theory of quantile functions / decreasing rearrangements:

Definition 3.2. Say $(\mathscr{G}, d_{\mathscr{G}})$ is a strongly left-invariant metric semigroup, and $X : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_{\mathscr{G}})$. The *decreasing* (or *non-increasing*) *rearrangement* of X is the right-continuous inverse X^* of the function $t \mapsto \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_0, z_0 X) > t \right)$, for any $z_0 \in \mathscr{G}$. In other words, X^* is the real-valued random variable defined on $[0,1]$ with the Lebesgue measure, as follows:

$$X^*(t) := \sup\{ y \in [0, \infty) : \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_0, z_0 X) > y \right) > t \}.$$

Now the proof of Theorem 3.1 uses the first assertion in [21, Proposition 6.8], extended here from Banach spaces to strongly left-invariant metric semigroups. It shows that controlling sums of $\mathscr{G}$-valued L^p random variables in probability (i.e., in L^0) allows us to control these sums in L^p as well, for $p > 0$.

Proposition 3.3. *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable strongly left-invariant metric semigroup, $p \in (0, \infty)$, and we have independent random variables $X_1, \dots, X_n \in L^p(\Omega, \mathscr{G})$, i.e., $\mathbb{E}_\mu[d_{\mathscr{G}}(z_0, z_0 X_j)^p] < \infty$ for all j (and any choice of $z_0 \in \mathscr{G}$). Now fix $z_0, z_1 \in \mathscr{G}$ and let S_k, U_n, M_n be as in (2.1) and Theorems 2.5, 2.6. Then,*

$$\mathbb{E}_\mu[U_n^p] \leqslant 2^{1+2p} (\mathbb{E}_\mu[M_n^p] + U_n^*(2^{-1-2p})^p).$$

As this result was not proved in [21], and given that we work here in significantly greater generality, we write down a proof for completeness.

Proof. The proof uses the following fact (which follows from the Pigeonhole principle):

$$0 \leqslant a \leqslant \sum_{j=1}^{k} w_j a_j,\ 0 \leqslant a_j, w_j\ \forall j \quad \Longrightarrow \quad a^p \leqslant (\sum_j w_j)^p \sum_{j=1}^{k} a_j^p\ \ \forall p \in (0,\infty). \tag{3.1}$$

Now to prove the result, note that $d_{\mathscr{G}}(z_1, z_0 S_k) \leqslant d_{\mathscr{G}}(z_1, z_0) + \sum_{j=1}^{k} d_{\mathscr{G}}(z_0, z_0 X_j)$ for all $1 \leqslant k \leqslant n$, by the triangle inequality and the strong left-invariance of $d_{\mathscr{G}}$. Using (3.1) with all $w_j = 1$,

$$U_n^p \leqslant \left(d_{\mathscr{G}}(z_1, z_0) + \sum_{j=1}^{n} d_{\mathscr{G}}(z_0, z_0 X_j) \right)^p \leqslant (n+1)^p \left(d_{\mathscr{G}}(z_1, z_0)^p + \sum_{j=1}^{n} d_{\mathscr{G}}(z_0, z_0 X_j)^p \right).$$

Hence $U_n \in L^p(\Omega, \mathscr{G})$ if all $X_j \in L^p(\Omega, \mathscr{G})$. Now recall that if $Z : (\Omega, \mathscr{A}, \mu) \to [0,\infty)$, then

$$\mathbb{E}_\mu[Z^\alpha] = \alpha \int_0^\infty t^{\alpha-1} \mathbb{P}_\mu\left(Z > t\right)\ dt, \qquad \forall \alpha > 0. \tag{3.2}$$

Apply Equation (3.2) with $Z = U_n$, $\alpha = p$, and $t \rightsquigarrow 4t$. Thus the integral converges, and we compute for any fixed $u \geqslant 0$ (using the Hoffmann-Jørgensen inequality in Theorem 2.6 with $k = 2, n_1 = n_2 = 1, t_1 = t_2 = s = t$, so that $I_0 = \{1,2\}$):

$$\begin{aligned}
\mathbb{E}_\mu[U_n^p] &= p \int_0^\infty (4t)^{p-1} \mathbb{P}_\mu\left(U_n > 4t\right)\ d(4t) = p4^p \int_0^\infty t^{p-1} \mathbb{P}_\mu\left(U_n > 4t\right)\ dt \\
&= p4^p \left(\int_0^u + \int_u^\infty \right) t^{p-1} \mathbb{P}_\mu\left(U_n > 4t\right)\ dt \\
&\leqslant 4^p u^p + p4^p \int_u^\infty t^{p-1} [\mathbb{P}_\mu\left(U_n > t\right)^2 + \mathbb{P}_\mu\left(M_n > t\right)]\ dt \\
&\leqslant 4^p u^p + p4^p \int_u^\infty t^{p-1} \mathbb{P}_\mu\left(U_n > t\right)^2\ dt + 4^p \mathbb{E}_\mu[M_n^p],
\end{aligned}$$

where the final inequality follows from (3.2) with $Z = M_n \in L^p(\Omega, \mathbb{R})$, $\alpha = p$. Now suppose $u > U_n^*(2^{-1-2p})$; then the outstanding integrand can be bounded above via

$$\mathbb{P}_\mu\left(U_n > t\right)^2 \leqslant \mathbb{P}_\mu\left(U_n > u\right) \mathbb{P}_\mu\left(U_n > t\right) \leqslant 2^{-1-2p} \mathbb{P}_\mu\left(U_n > t\right).$$

Continuing with the above calculations,

$$\mathbb{E}_\mu[U_n^p] \leqslant (4u)^p + 4^p \mathbb{E}_\mu[M_n^p] + 4^p \cdot 2^{-1-2p} \cdot \int_0^\infty pt^{p-1} \mathbb{P}_\mu\left(U_n > t\right)\ dt = (4u)^p + 4^p \mathbb{E}_\mu[M_n^p] + 2^{-1} \mathbb{E}_\mu[U_n^p],$$

by a third application of (3.2) with $Z = U_n$, $\alpha = p$. Since this inequality holds for all $u > U_n^*(2^{-1-2p})$, the desired claim follows. □

With Proposition 3.3 at hand, one shows Theorem 3.1 by closely following the proof of [18, Theorem A]. Here we only point out the modifications required: in addition to using Proposition 3.3 over strongly left-invariant metric semigroups, the only other update is that in showing that

$$d_{\mathscr{G}}(z_0, z_0 X_n) \leqslant d_{\mathscr{G}}(z_1, z_0 S_{n-1}) + d_{\mathscr{G}}(z_1, z_0 S_n),$$

we now use the strong left-invariance rather than the bi-invariance of the metric $d_{\mathscr{G}}$. □

The other result that we discuss and extend here (from bi-invariant to strongly left-invariant metric semigroups), again uses the Hoffmann-Jørgensen inequality to relate the L^p-norm of U_n to its tail distribution (using U_n^*).

Proposition 3.4. *There exists a universal positive constant c_1 such that for any $0 \leqslant t \leqslant s \leqslant 1/2$, any separable strongly left-invariant metric semigroup $(\mathscr{G}, d_{\mathscr{G}})$ with elements z_0, z_1, and any sequence of independent $\mathscr{G}$-valued random variables $X_1, \ldots, X_n$,*

$$U_n^*(t) \leqslant c_1 \frac{\log(1/t)}{\max\{\log(1/s), \log\log(4/t)\}} (U_n^*(s) + M_n^*(t/2)),$$

with $U_n := \max_{1 \leqslant j \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_j)$ and $M_n := \max_{1 \leqslant j \leqslant n} d_{\mathscr{G}}(z_0, z_0 X_j)$ as above.

We again omit the proof, as it is the same as that of [18, Lemma 3.1], via the proof of [9, Corollary 1]. □

As a final application to conclude this section: in [1], Bagyan–Richards applied the results in [17] to derive similar results and tail bounds (and more) for random walks on the cone/Riemannian manifold of positive definite matrices – more precisely, for products of such matrices, randomly chosen according to an orthogonally bi-invariant distribution. Given the extension in this work of the Hoffmann-Jørgensen inequality to strongly left-invariant semigroups (equivalently, to left-invariant monoids, as is shown below) – including all Lie groups – we expect that a broader range of applications of the kind in [1] can now be deduced in random matrix theory.

4. Additional probability inequalities and Lévy's equivalence

We next extend several other probability inequalities to strongly left/right-invariant metric semigroups (in particular, they now become applicable to all Lie groups), from their previous avatars over either bi-invariant metric semigroups, or in one case, Banach spaces. These inequalities also help extend the following result:

Theorem 4.1 (Lévy's equivalence)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a complete separable strongly left-invariant metric semigroup, $X_n : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_{\mathscr{G}})$ are independent, $X \in L^0(\Omega, \mathscr{G})$, and S_n is as in* (2.1)*. Then*

$$S_n \longrightarrow X \text{ a.s. } \mathbb{P}_\mu \quad \Longleftrightarrow \quad S_n \xrightarrow{P} X.$$

If instead the sequence S_n does not converge in the above manner, then it diverges almost surely.

Via Proposition 2.10, this result simultaneously extends [18, Theorem 2.1] for $\mathscr{G}$ a complete separable bi-invariant metric semigroup – which in turn had extended variants for $\mathscr{G}$ a separable Banach space by Itô–Nisio [13, Theorem 3.1] and by Hoffmann-Jørgensen and Pisier [12, Lemma 1.2] – as well as Tortrat's result in [28] for $\mathscr{G}$ a complete separable metric group.

Remark 4.2. We claim that the setting in Theorem 4.1 is strictly more general than even [18, Theorem 2.1] – which itself was more general than the preceding variants. This follows by considering examples of Lie groups with left-invariant but not bi-invariant Riemannian metric (by the results in [24], as discussed above).

The proof of Theorem 4.1 employs the Ottaviani–Skorohod inequality:

Proposition 4.3 (Ottaviani–Skorohod inequality)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable strongly left-invariant metric semigroup, and $X_1, \ldots, X_n : \Omega \to \mathscr{G}$ are independent. Fix $0 < \alpha, \beta \in \mathbb{R}$. Then for all $z_0, z_1 \in \mathscr{G}$,*

$$\mathbb{P}_\mu \left(\max_{1 \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k) \geqslant \alpha + \beta \right) \cdot \min_{1 \leqslant k \leqslant n} \mathbb{P}_\mu \left(d_{\mathscr{G}}(S_k, S_n) \leqslant \beta \right) \leqslant \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_1, z_0 S_n) \geqslant \alpha \right).$$

This is a special case of the Mogul'skii inequalities (setting $m = 1, a = \alpha + \beta, b = \beta$ in the next result) – in [18] for bi-invariant metric semigroups, but also more generally in the present setting:

Proposition 4.4 (Mogul'skii inequalities)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable strongly left-invariant metric semigroup, $z_0, z_1 \in \mathscr{G}$, $a, b, c \in [0, \infty)$, and $X_1, \ldots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. If $1 \leqslant m \leqslant n$ in $\mathbb{Z}$, then:*

$$\mathbb{P}_\mu\left(\min_{m \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k) \leqslant a\right) \cdot \min_{m \leqslant k \leqslant n} \mathbb{P}_\mu\left(d_{\mathscr{G}}(S_k, S_n) \leqslant b\right) \leqslant \mathbb{P}_\mu\left(d_{\mathscr{G}}(z_1, z_0 S_n) \leqslant a + b\right),$$

$$\mathbb{P}_\mu\left(\max_{m \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k) \geqslant a\right) \cdot \min_{m \leqslant k \leqslant n} \mathbb{P}_\mu\left(d_{\mathscr{G}}(S_k, S_n) \leqslant b\right) \leqslant \mathbb{P}_\mu\left(d_{\mathscr{G}}(z_1, z_0 S_n) \geqslant a - b\right).$$

It is the formulation of the above results in this greater generality that is of note; we omit all three of the proofs here (as in the preceding section), as they go through verbatim except for a common workaround needed in one step in each proof. Namely: one needs to equate $d_{\mathscr{G}}(z_0 S_k, z_0 S_n)$ for $k \leqslant n$ with $d_{\mathscr{G}}(z_0, z_0 X_{k+1} \cdots X_n)$ – and this follows from the strong left-invariance of $\mathscr{G}$. □

The next result is the Lévy–Ottaviani inequality, now extended from reals to our general setting:

Proposition 4.5 (Lévy–Ottaviani inequality)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable strongly left-invariant metric semigroup, $z_0, z_1 \in \mathscr{G}$, and $X_1, \ldots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. Given $a \geqslant 0$, define:*

$$U_n := \max_{1 \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k), \qquad p_a := \max_{1 \leqslant k \leqslant n} \mathbb{P}_\mu\left(d_{\mathscr{G}}(z_1, z_0 S_k) > a\right),$$

where $S_k := X_1 \cdots X_k$ for all k, as above. Then for all $l \geqslant 2$ and $a_1, \ldots, a_l \geqslant 0$,

$$\mathbb{P}_\mu\left(U_n > a_1 + \cdots + a_l\right) \leqslant \sum_{i=2}^{l} p_{a_i} + p'_l,$$

where $p'_l := p_{a_1}$ if l is odd, and $p'_l := \max_{1 \leqslant k \leqslant n} \mathbb{P}_\mu\left((d_{\mathscr{G}}(S_k, S_n) > a_1\right)$ if l is even.

This result generalizes [3, Theorem 22.5], both in its setting and its statement. To see this, set

$$l = 3, \qquad a_1 = a_2 = a_3 = \alpha, \qquad \mathscr{G} = (\mathbb{R}, +), \qquad z_0 = z_1 = 0.$$

Also note that the result is false if $l = 1$ and $p'_1 = p_{a_1}$, since $\mathbb{P}_\mu\left(d_{\mathscr{G}}(z_1, z_0 S_k) > a_1\right) \leqslant \mathbb{P}_\mu\left(U_n > a_1\right)$ for all k.

Proof. We write down a proof in this general setting, to indicate where one needs to use strong left-invariance (as opposed to any group or Banach space structure). Define the stopping time $\tau : \Omega \to \{1, \ldots, n\}$ via:

$$\tau = \inf\{k \in [1, n] \cap \mathbb{Z} \ : \ d_{\mathscr{G}}(z_1, z_0 S_k) > a_1 + \cdots + a_l\}.$$

It is now not hard to show that:

$$\begin{aligned}
\mathbb{P}_\mu\left(U_n > \sum_{i=1}^{l} a_i\right) &\leqslant \mathbb{P}_\mu\left(d_{\mathscr{G}}(z_1, z_0 S_n) > a_l\right) + \sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k, d_{\mathscr{G}}(z_1, z_0 S_n) \leqslant a_l\right) \\
&\leqslant p_{a_l} + \sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k, d_{\mathscr{G}}(S_k, S_n) > \sum_{i=1}^{l-1} a_i\right) \\
&= p_{a_l} + \sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k\right) \mathbb{P}_\mu\left(d_{\mathscr{G}}(S_k, S_n) > \sum_{i=1}^{l-1} a_i\right),
\end{aligned}$$

where the final equality uses the independence of $\tau = k$ from the behavior of $d_{\mathscr{G}}(S_k, S_n) = d_{\mathscr{G}}(z_0, z_0 X_{k+1} \cdots X_n)$; this last uses the strong left-invariance of $d_{\mathscr{G}}$. There are now two cases:

- If $l = 2$, then the last probability (inside the sum) is bounded above by p'_l for $l = 2$, and the result follows since $\sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k\right) \leqslant 1$.

- If $l > 2$ and $d_{\mathscr{G}}(S_k, S_n) > \sum_{i=1}^{l-1} a_i$, then either $d_{\mathscr{G}}(z_1, z_0 S_k) > a_{l-1}$, or $d_{\mathscr{G}}(z_1, z_0 S_n) > \sum_{i=1}^{l-2} a_i$. By definition, the first event has probability at most $p_{a_{l-1}}$, while the probability of the latter event is analyzed in two sub-cases: first, if $l = 3$, then it is dominated by $p'_l = p_{a_1}$, which proves the result as follows:

$$\mathbb{P}_\mu\left(U_n > \sum_{i=1}^{3} a_i\right) \leqslant p_{a_3} + \sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k\right)\left(p_{a_2} + p_{a_1}\right) \leqslant p_{a_3} + 1 \cdot (p_{a_2} + p'_l).$$

Next, if $l > 3$, then we prove the result by using induction and the above analysis for the base cases of $l = 2, 3$: the second event is dominated by $U_n > \sum_{i=1}^{l-2} a_i$. Hence by the induction hypothesis for $l - 2$,

$$\begin{aligned}\mathbb{P}_\mu\left(U_n > \sum_{i=1}^{l} a_i\right) &\leqslant p_{a_l} + \sum_{k=1}^{n-1} \mathbb{P}_\mu\left(\tau = k\right)\left(p_{a_{l-1}} + \mathbb{P}_\mu\left(U_n > \sum_{i=1}^{l-2} a_i\right)\right) \\ &\leqslant p_{a_l} + 1 \cdot \left(p_{a_{l-1}} + \sum_{i=2}^{l-2} p_{a_i} + p'_{l-2}\right),\end{aligned}$$

and we are done since p'_l depends only on the parity of l. □

Remark 4.6. We quickly discuss additional results in the present, general setting. First, several other results in the literature were shown in [17, 18, 19] for random variables valued in abelian metric semigroups $\mathscr{G}_{ab}$. These results automatically extend to the present setting (since the metric in $\mathscr{G}_{ab}$ is bi-invariant, hence strongly left-invariant).

Second, further results were shown in *loc. cit.*, over all semigroups with a bi-invariant metric. Given the above results in the paper, we expect that these latter results should hold for all strongly left/right-invariant metric semigroups – for instance, the results on the Lévy property found in [9, §4]. We do not pursue this investigation in the present work, leaving it to the interested reader.

Third, we leave the parallel formulation and proofs of the strongly right-invariant analogues of these results (discussed here) and of the ones studied in this section, to the interested reader.

5. Embedding (strong) left-invariant semigroups in monoids, following Malcev

This concluding section is not concerned with results in probability, but with the framework underlying this entire work: strongly left-invariant metric semigroups. In previous work and in the results above, we have seen random variables take values in three different competing structures (which are each more primitive than normed linear spaces, or even groups):

(i) bi-invariant metric semigroups/monoids,
(ii) (strongly) left-invariant metric monoids,
(iii) strongly left-invariant metric semigroups.

By Proposition 2.10, every bi-invariant metric semigroup embeds isometrically and homomorphically into a left-invariant (in fact, bi-invariant) metric monoid – which is automatically a strongly left- and right-invariant monoid. Thus, (i) is a strictly more restrictive notion than (ii), since – as explained above using [24] – there exist Lie groups that are equipped with (strongly) left-invariant metrics that are not bi-invariant.

The other comparison question is between strongly left-invariant metric (ii) monoids and (iii) semigroups (or right-invariant). If indeed these notions are the same, then this entire paper could have equivalently been written for left-invariant metric monoid-valued random variables. And indeed, we now show this to be the case:

Theorem 5.1. *Let $\mathscr{G}$ be a semigroup with a left-invariant metric $d_{\mathscr{G}}$.*

(1) *Every idempotent in $\mathscr{G}$ is a left-identity. Thus, every right-identity is a two-sided identity. (The converse statements are obvious.)*
(2) *Suppose $\mathscr{G}$ has no two-sided/right identity. Then $\mathscr{G}$ embeds isometrically and homomorphically inside a left-invariant metric monoid (onto the non-identity elements) if and only if $d_{\mathscr{G}}$ is strongly left-invariant. In particular, such a semigroup $\mathscr{G}$ has no idempotents.*

We leave it to the interested reader to formulate and prove the (obvious) counterpart for right-invariant metric semigroups $\mathscr{G}$. Also, this situation is parallel to the case of bi-invariant metric semigroups, which always embed inside a bi-invariant metric monoid by Proposition 2.10.

Proof. First, if $e^2 = e \in \mathscr{G}$ then $d_{\mathscr{G}}(eg, g) = d_{\mathscr{G}}(e^2 g, eg) = 0$, so $eg = g$ for all $g \in \mathscr{G}$, as claimed. In particular, since a right-identity is an idempotent, it is automatically a two-sided identity.

Next, if $\mathscr{G}$ embeds isometrically (and strictly) inside a left-invariant metric monoid, say $(\mathscr{G}_\circ, e, d)$, then $\mathscr{G}$ is strongly left-invariant (as repeatedly seen above). Moreover, $\mathscr{G}$ has no idempotent, for if g is any idempotent in $\mathscr{G}_\circ$ then $d(e, g) = d(g, g^2) = 0$, so $g = e \notin \mathscr{G}$.

Conversely, if $d_{\mathscr{G}}$ is strongly left-invariant, define $\mathscr{G}_\circ := \mathscr{G} \sqcup \{e\}$, and extend the semigroup product to $\mathscr{G}_\circ$ via: $e \cdot g = g \cdot e := g$ for all $g \in \mathscr{G}_\circ$. Also extend the metric $d_{\mathscr{G}}$ to a symmetric map d on all of $\mathscr{G}_\circ \times \mathscr{G}_\circ$ via: $d \equiv d_{\mathscr{G}}$ on $\mathscr{G} \times \mathscr{G}$, $d(e, e) := 0$, and

$$d(e, g) = d(g, e) := d_{\mathscr{G}}(g, g^2) \ \forall g \in \mathscr{G}.$$

We claim this last expression is positive. If not, then $g^2 = g$, so g is a left-identity by the preceding part. Moreover, by strong left-invariance,

$$d_{\mathscr{G}}(a, ag) = d_{\mathscr{G}}(g, g^2) = 0 \ \forall a \in \mathscr{G},$$

so g is a two-sided identity, which is false. We next claim that d satisfies the triangle inequality – for which it suffices to work with $a, b \in \mathscr{G}$ and e. This too is verified using strong left-invariance:

$$d(a, e) + d(e, b) = d_{\mathscr{G}}(a^2, a) + d_{\mathscr{G}}(b, b^2) = d_{\mathscr{G}}(ba, b) + d_{\mathscr{G}}(b, b^2) \geqslant d_{\mathscr{G}}(ba, b^2) = d(a, b),$$
$$d(a, b) + d(e, b) = d_{\mathscr{G}}(ba, b^2) + d_{\mathscr{G}}(b, b^2) \geqslant d_{\mathscr{G}}(ba, b) = d_{\mathscr{G}}(a^2, a) = d(a, e).$$

Hence d is a metric. That d is left-invariant follows from the strong left-invariance of $d_{\mathscr{G}}$. □

Following the comparisons between the primitive settings (i)–(iii) at the start of this section, a final comparison to make is between (ii)=(iii) (shown above) and (iv) left-invariant metric semigroups. Namely: *Can the word "strong" be removed from Theorem 5.1(2)?* While the proofs above use strong left-invariance, it is not clear if this itself follows (or not) from "usual" left-invariance in any semigroup, as it does in all monoids. Thus, if the question has a positive answer, this paper could even have been written over all (iv) left-invariant metric semigroups. However, we now show this is not true – even for finitely generated complete metric semigroups, cf. Theorem 4.1 – by providing a counterexample to the question above (also alluded to in the opening section):

Proposition 5.2. *There exists a countable, discrete (hence complete) left-invariant metric semigroup $\mathscr{G}$ with two generators, which is not a monoid but contains an idempotent.*

In particular, by Theorem 5.1 $\mathscr{G}$ can never embed into a left-invariant metric monoid, and $d_{\mathscr{G}}$ is not strongly left-invariant.

Proof. On the countable set $\mathscr{G} := \{h^{n+1}, h^n g : n \in \mathbb{Z}_{\geqslant 0}\}$, define the operation $h^n g^\varepsilon \cdot h^{n'} g^{\varepsilon'} := h^{n+n'} g^{\varepsilon'}$. Then $\mathscr{G}$ is a semigroup generated by g, h, with g a left-identity (and the only such), hence idempotent; but as $h^{n+1} g \neq h^{n+1}$ for each $n \geqslant 0$, $\mathscr{G}$ has no two-sided identity, so is not a monoid.

Next, define the Manhattan distance $d_{\mathscr{G}} : \mathscr{G} \times \mathscr{G} \to [0, \infty)$ via: $d_{\mathscr{G}}(h^n g^\varepsilon, h^{n'} g^{\varepsilon'}) := |n - n'| + |\varepsilon - \varepsilon'|$. Clearly, $d_{\mathscr{G}}$ is positive and symmetric. The triangle inequality is verified by checking six easy cases; we show one case here. For all $n' > 0$ and $n \geqslant 0$,

$$d_{\mathscr{G}}(h^{n'}, h^n g) + d_{\mathscr{G}}(h^{n'}, h^m) = |n' - n| + 1 + |n' - m| \geqslant |n - m| + 1 = d_{\mathscr{G}}(h^m, h^n g).$$

Thus $d_{\mathscr{G}}$ is a metric; it is easy to see that $(\mathscr{G}, d_{\mathscr{G}})$ is discrete, hence complete.

The next step is to check that $d_{\mathscr{G}}$ is left-invariant, but not strongly so. For the latter, we compute

$$d_{\mathscr{G}}(h^n g, h^n g \cdot h^{n'}) = n' + 1 \neq n' = d_{\mathscr{G}}(h^{n'}, h^{n'} \cdot h^{n'}), \quad \forall n' > 0,\ n \geqslant 0,$$

while the former is straightforward – as a sample calculation, we have

$$d_{\mathscr{G}}(h^n g^{\varepsilon} \cdot h^{n'} g^{\varepsilon'}, h^n g^{\varepsilon} \cdot g) = n' + 1 - \varepsilon' = d_{\mathscr{G}}(h^{n'} g^{\varepsilon'}, g)$$

for all $n' > 0$, $\varepsilon' \in \{0, 1\}$, and $h^n g^{\varepsilon} \in \mathscr{G}$ (i.e. $(\varepsilon, n) \neq (0, 0)$). □

Remark 5.3. Another remark, for completeness, is that given a left-invariant metric *group* $\mathscr{G}$, its metric $d_{\mathscr{G}}$ is 2-homogeneous – i.e., $d_{\mathscr{G}}(e, g^2) = 2d_{\mathscr{G}}(e, g)$ for all $g \in \mathscr{G}$ – if and only if $\mathscr{G}$ is abelian. (In particular, $d_{\mathscr{G}}$ is bi-invariant.) This follows from the recent Polymath project [27].

For completeness, and given the above discussion in this section, we conclude with a follow-up question to the above results: *Does every left-invariant metric monoid embed inside a left-invariant metric group?* The answer is in the negative for an even more general question: for bi-invariant metric monoids to always embed inside some group! Namely, if one takes $d_{\mathscr{G}}$ to be the discrete metric, then note that $d_{\mathscr{G}}$ is bi-invariant if and only if the monoid $\mathscr{G}$ is cancellative: $ac = bc$ or $ca = cb$ implies $a = b$. Now Malcev [23] has constructed a monoid with eight generators that is cancellative, but cannot map injectively and homomorphically into any group (metric or not).

Remark 5.4 (The idea of Malcev)**.** The preceding question occurred to us as a natural parallel to the question answered negatively in this section: *Is every left-invariant metric semigroup strongly left-invariant, i.e. does every left-invariant metric semigroup embed isometrically and homomorphically into a left-invariant metric monoid?* Moreover, we learned of Malcev's (negative) example shortly after having shown the results in this paper. That said, not only are the two questions similar, and their answers both negative, but interestingly, even the approaches have a common philosophy. Namely, our approach in Theorem 5.1 and Proposition 5.2 was to first show that the only idempotent in a strongly left-invariant semigroup is a (=the) two-sided identity; and to then construct a left-invariant semigroup containing an idempotent that is not a two-sided identity. This is precisely the philosophy behind Malcev's beautiful and more intricate idea in [23], which was to first note that if a group contains eight elements a, b, c, d, u, v, x, y that satisfy the relations

$$au = bv, \qquad ax = by, \qquad cx = dy, \tag{5.1}$$

then one can successively solve to get $vu^{-1} = b^{-1}a = yx^{-1} = d^{-1}c$, and hence $cu = dv$. In the second step, Malcev constructed a monoid M that is cancellative (note, this is if and only if the discrete metric on M is bi-invariant), and is finitely generated with generators a, b, c, d, u, v, x, y that satisfy (5.1), but in which $cu \neq dv$. Hence M cannot embed as a sub-monoid in any group.

Acknowledgements. I thank Terence Tao for the reference [23], and Bhaswar Bhattacharya, Manjunath Krishnapur, Muna Naik, and Soumik Pal for useful discussions. This work was partially supported by Ramanujan Fellowship grant SB/S2/RJN-121/2017 and SwarnaJayanti Fellowship grants SB/SJF/2019-20/14 and DST/SJF/MS/2019/3 from SERB and DST (Govt. of India), by a Shanti Swarup Bhatnagar Award from CSIR (Govt. of India), and by the DST FIST program 2021 [TPN–700661].

References

[1] A. Bagyan and D. St. P. Richards, *Hoffmann-Jørgensen inequalities for random walks on the cone of positive definite matrices*, Journal of Theoretical Probability 36 no. 2, 1181–1202, 2023.

[2] Y. Benoist and J.-F. Quint, *Central limit theorem for linear groups*, Annals of Probability 44 no. 2, 1308–1340, 2016.

[3] P. Billingsley, *Probability and Measure*, Wiley series in Probability & Mathematical Statistics, Wiley and Sons, New York, 1995.

[4] P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability Theory and Related Fields) **57**, 159–179, 1981.
[5] H. Furstenberg and H. Kesten, *Products of random matrices*, Annals of Mathematical Statistics **31** no. 2, 457–469, 1960.
[6] E. Giné and J. Zinn, *Central Limit Theorems and Weak Laws of Large Numbers in certain Banach spaces*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability Theory and Related Fields) **62** no. 3, 323–354, 1983.
[7] U. Grenander, *Probabilities on Algebraic Structures*, Wiley, New York, 1963.
[8] A. Guionnet, M. Krishnapur, and O. Zeitouni, *The single ring theorem*, Annals of Mathematics (2) 174 no. 2, 1189–1217, 2011.
[9] P. Hitczenko and S.J. Montgomery-Smith, *Measuring the magnitude of sums of independent random variables*, Annals of Probability **29**, 447–466, 2001.
[10] M. Hochman and B. Solomyak, *On the dimension of Furstenberg measure for $SL_2(\mathbb{R})$ random matrix products*, Inventiones Mathematicae **210** no. 3, 815–875, 2017.
[11] J. Hoffmann-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Mathematica **52**, 159–186, 1974.
[12] J. Hoffmann-Jørgensen and G. Pisier, *The Law of Large Numbers and the Central Limit Theorem in Banach spaces*, Annals of Probability **4**, 587–599, 1976.
[13] K. Itô and M. Nisio, *On the convergence of sums of independent Banach space valued random variables*, Osaka Journal of Mathematics **5**, 35–48, 1968.
[14] K. Johansson, *On random matrices from the compact classical groups*, Annals of Mathematics (2) 145 no. 3, 519–545, 1997.
[15] W.B. Johnson and G. Schechtman, *Sums of independent random variables in rearrangement invariant function spaces*, Annals of Probability **17**, 789–808, 1989.
[16] J.-P. Kahane, *Some Random Series of Functions*. Vol. **5** in Cambridge Studies in Advanced Mathematics (2nd ed.), Cambridge Univ. Press, London, 1968.
[17] A. Khare and B. Rajaratnam, *The Hoffmann-Jørgensen inequality in metric semigroups*, Annals of Probability **45** no. 6A, 4101–4111, 2017.
[18] A. Khare and B. Rajaratnam, *Probability inequalities and tail estimates for metric semigroups*, Advances in Operator Theory **5** no. 3, 779–795, 2020.
[19] A. Khare and B. Rajaratnam, *The Khinchin–Kahane and Lévy inequalities for abelian metric groups, and transfer from normed (abelian semi)groups to Banach spaces*, Journal of Mathematical Analysis and Applications **528** no. 2, art.# 127545 (18 pp.), 2023.
[20] M.J. Klass and K. Nowicki, *An improvement of Hoffmann-Jørgensen's inequality*, Annals of Probability **28**, 851–862, 2000.
[21] M. Ledoux and M. Talagrand, *Probability in Banach Spaces (Isoperimetry and Processes)*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, Berlin-New York, 1991.
[22] L. Lovász, *Large Networks and Graph Limits*, Vol. **60** in *Colloquium Publications*, American Mathematical Society, Providence, 2012.
[23] A. Malcev. *On the immersion of an algebraic ring into a field*, Mathematische Annalen **113**, 686–691, 1937.
[24] J. Milnor, *Curvatures of left invariant metrics on Lie groups*, Advances in Mathematics **21** no. 3, 293–329, 1976.
[25] K.R. Parthasarathy, *Probability measures on metric spaces*, Probability and Mathematical Statistics series of Monographs and Textbooks, Academic Press, New York, 1967.
[26] M. Pollicott, *Maximal Lyapunov exponents for random matrix products*, Inventiones Mathematicae **181** no. 1, 209–226, 2010.
[27] D.H.J. Polymath (T. Fritz, S. Gadgil, A. Khare, P. Nielsen, L. Silberman, and T. Tao), *Homogeneous length functions on groups*, Algebra & Number Theory **12** no. 7, 1773–1786, 2018.
[28] A. Tortrat, *Lois de probabilite sur un espace topologique completement régulier et produits infinis à termes indépendant dans un groupe topologique*, Annales de l'Institut Henri Poincaré (B): Probabilités et Statistique **1**, 217–237, 1964/65.

(A. Khare) Indian Institute of Science, Bangalore – 560012, India; and Analysis and Probability Research Group, Bangalore – 560012, India
Email address: khare@iisc.ac.in