The Annals of Probability
2017, Vol. 45, No. 6A, 4101–4111
DOI: 10.1214/16-AOP1160

THE HOFFMANN–JØRGENSEN INEQUALITY IN METRIC SEMIGROUPS[1]

BY APOORVA KHARE* AND BALA RAJARATNAM*,†,‡

*Stanford University**, *University of California, Davis*†, *and University of Sydney*‡

We prove a refinement of the inequality by Hoffmann–Jørgensen that is significant for three reasons. First, our result improves on the state-of-the-art even for real-valued random variables. Second, the result unifies several versions in the Banach space literature, including those by Johnson and Schechtman [*Ann. Probab.* **17** (1989) 789–808], Klass and Nowicki [*Ann. Probab.* **28** (2000) 851–862], and Hitczenko and Montgomery-Smith [*Ann. Probab.* **29** (2001) 447–466]. Finally, we show that the Hoffmann–Jørgensen inequality (including our generalized version) holds not only in Banach spaces but more generally, in a very primitive mathematical framework required to state the inequality: a metric semigroup $\mathscr{G}$. This includes normed linear spaces as well as all compact, discrete or (connected) abelian Lie groups.

1. Introduction. In this paper, our goal is to present a broad generalization of the Hoffmann–Jørgensen inequality (see Theorem A). This is a classical result in the literature, which is widely used in bounding sums of independent random variables, with several different versions proved in the general setting of a separable Banach space (see [3, 5, 11, 13]). We recall a "first version" from the literature.

THEOREM 1 (Ledoux and Talagrand, [13], Proposition 6.7). *Suppose $\mathbb{B}$ is a separable Banach space, and $(\Omega, \mathscr{A}, \mu)$ is a probability space with $X_1, \dots, X_n \in L^0(\Omega, \mathbb{B})$ independent random variables. For $1 \le j \le n$, define $S_j := X_1 + \cdots + X_j$ and $U_n := \max_{1\le j\le n} \|S_j\|$. Then*

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

This version incorporates results by Kahane [6] and Hoffmann–Jørgensen [4]. (See also [3] for a detailed history of the inequality.)

Theorem 1 has seen subsequent generalizations by several authors, including Johnson and Schechtman [5], Klass and Nowicki [11], and Hitczenko and Montgomery-Smith [3]. This last variant is now stated as the following.

Received May 2016; revised October 2016.

[1]Supported in part by US Air Force Office of Scientific Research Grant award FA9550-13-1-0043, US National Science Foundation under Grant DMS-0906392, DMS-CMG 1025465, AGS-1003823, DMS-1106642, DMS-CAREER-1352656, Defense Advanced Research Projects Agency DARPA YFA N66001-111-4131, the UPS Foundation and SMC-DBNKY.

MSC2010 subject classifications. Primary 60E15; secondary 60B15.

Key words and phrases. Hoffmann–Jørgensen inequality, metric semigroup.

THEOREM 2 (Hitczenko and Montgomery-Smith, [3], Theorem 1). (*Notation as in Theorem* 1.) *For all* $K \in \mathbb{N}$ *and* $s, t \in (0, \infty)$,

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

While isoperimetric methods provide more powerful techniques to work with, the aforementioned manifestations of the Hoffmann–Jørgensen inequality for Banach spaces also have numerous consequences in estimating the magnitude and behavior of the quantities $\|S_n\|$ and U_n, as explained in [3, 11], for instance.

We now present several motivations behind the present paper. First, our main result in Theorem A provides an improvement on Theorems 1 and 2 above. Note, Theorem 2 has a variant via the order statistics of the variables $Y_j := \|X_j\|$ (see [3]). Our result improves on this strengthening as well.

Second, it is not clear if either of Theorems 1 or 2 follows from the other, or if they are even logically related. Our result (Theorem A) simultaneously unifies and significantly generalizes both of these results.

A third motivation arises out of independent mathematical and applied interest. Note that to state the above inequalities, one requires merely the notions of a metric and a binary associative operation. Thus, a question of interest is to ascertain whether the result holds in the more general setting of a separable metric semigroup $\mathscr{G}$ (defined below).

In this paper, we provide a positive answer to the above question. Thus, we show Theorem A in a very primitive mathematical setting required to state the Hoffmann–Jørgensen inequality. Our motivations in so doing are both modern as well as traditional. Classically, a cornerstone of twentieth-century probability theory has been the systematic and rigorous development of the field, for random variables taking values in Banach spaces. At the same time, general results on Fourier analysis and Haar measure for compact abelian groups, and the study of random variables with values in metric groups [2, 15] motivate the need to develop results in the greatest possible generality. The present paper lies squarely in this area.

Additionally, an increasing number of modern-day settings involve working outside the traditional Banach space paradigm. Indeed, settings of compact and abelian Lie groups are studied in the literature, including permutation groups, lattices and other discrete (semi)groups, circle groups and tori. Moreover, modern data are manifold-valued—including in real/complex Lie groups—as opposed to the traditionally well-studied normed linear spaces. Other modern settings include the space of graphons with the cut-norm [14], as well as the space of labelled graphs $\mathscr{G}(V)$ on a fixed vertex set V, which was studied in [7, 8]. The space $\mathscr{G}(V)$ turns out to be a 2-torsion group, and hence cannot embed as a subgroup into a normed linear space. Thus, Banach space methods are not adequate to study stochastic phenomena in modern-day settings. To this end, this paper allows for studying tail estimates and bounding random sums in greater generality.

2. Metric semigroups and the main result. We now set some notation and state our main result.

DEFINITION 3. A *metric semigroup* is defined to be a semigroup $(\mathscr{G}, \cdot)$ equipped with a metric $d_{\mathscr{G}} : \mathscr{G} \times \mathscr{G} \to [0, \infty)$ that is translation-invariant:

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

Equivalently, $(\mathscr{G}, d_{\mathscr{G}})$ is a metric space equipped with a associative binary operation $\cdot$ such that $d_{\mathscr{G}}$ is translation-invariant.

Metric (semi)groups are ubiquitous in probability theory. Examples include Banach spaces such as function spaces, discrete semigroups (including finite groups as well as labelled graph space $\mathscr{G}(V)$ [7, 8]), and all compact or (connected) abelian Lie groups, which include the circle and tori (via, e.g., [16], Theorem V.5.3). Among other examples are amenable groups (see [1], Proposition 4.12, and the discussion around it) and abelian Hausdorff metrizable topologically complete groups [12].

DEFINITION 4. Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable metric semigroup, with Borel σ-algebra $\mathscr{B}_{\mathscr{G}}$. Given integers $1 \leq j \leq n$ and random variables $X_1, \dots, X_n : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_{\mathscr{G}})$, define

$$S_j(\omega) := X_1(\omega) \cdots X_j(\omega), \qquad M_j(\omega) := \max_{1 \leq i \leq j} d_{\mathscr{G}}\big(z_0, z_0 X_i(\omega)\big), \tag{2}$$

where $z_0 \in \mathscr{G}$ is arbitrary. (We show below, M_j is independent of $z_0 \in \mathscr{G}$.)

We now state our main result, namely, the aforementioned generalization of the Hoffmann–Jørgensen inequality, for separable metric semigroups.

THEOREM A. *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable metric semigroup, $z_0, z_1 \in \mathscr{G}$ are fixed, and $X_1, \dots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. Also fix integers $k, n_1, \dots, n_k \in \mathbb{N}$ and nonnegative scalars $t_1, \dots, t_k, s \in [0, \infty)$, and define*

$$U_n := \max_{1 \leq j \leq n} d_{\mathscr{G}}(z_1, z_0 S_j), \qquad I_0 := \left\{1 \leq i \leq k : \mathbb{P}_\mu(U_n \leq t_i)^{n_i - \delta_{i1}} \leq \frac{1}{n_i!}\right\}, \tag{3}$$

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

$$\begin{aligned} &\mathbb{P}_\mu\left(U_n > (2n_1 - 1)t_1 + 2\sum_{i=2}^k n_i t_i + \left(\sum_{i=1}^k n_i - 1\right)s\right) \\ &\quad \leq \mathbb{P}_\mu(U_n \leq t_1)^{\mathbf{1}_{1 \notin I_0}} \prod_{i \in I_0} \mathbb{P}_\mu(U_n > t_i)^{n_i} \prod_{i \notin I_0} \frac{1}{n_i!} \left(\frac{\mathbb{P}_\mu(U_n > t_i)}{\mathbb{P}_\mu(U_n \leq t_i)}\right)^{n_i} \\ &\qquad + \mathbb{P}_\mu(M_n > s). \end{aligned} \tag{4}$$

More generally, define

$$K := \sum_{i=1}^{k} n_i, \qquad Y_j := d_{\mathscr{G}}(z_0, z_0 X_j),$$

$$Y_{(1)} := \min(Y_1, \dots, Y_n), \dots, Y_{(n)} := \max(Y_1, \dots, Y_n),$$

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

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

Theorem A generalizes the original Hoffmann–Jørgensen inequality in many ways: mathematically it is a significant generalization of Theorem 1 (which itself generalizes the classical Hoffmann–Jørgensen inequality for Euclidean, Hilbert and Banach spaces). To see this, set

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

Now Theorem 1 follows from Theorem A with $I_0 = \{1, 2\}$.

Moreover, Theorem A also generalizes [3], Theorem 1, that is, Theorem 2—which has different bounds than Theorem 1. To see this, set $\mathscr{G} = \mathbb{B}, z_0 = z_1 = 0, k = 1, n_1 = K, t_1 = t$. Now the first expression on the right-hand side of equation (4) can be rewritten as follows:

$$\prod_{i=1}^{k} \mathbb{P}_\mu(U_n > t_i)^{n_i} \min\left(1, \frac{1}{n_i! \cdot \mathbb{P}_\mu(U_n \leq t_i)^{n_i - \delta_{i1}}}\right). \tag{5}$$

Thus, with the above values, Theorem 2 follows from Theorem A:

$$\begin{aligned}
&\mathbb{P}_\mu\bigl(U_n > 2Kt + (K-1)s\bigr) \\
&\quad \leq \mathbb{P}_\mu\bigl(U_n > (2K-1)t + (K-1)s\bigr) \\
&\quad \leq \mathbb{P}_\mu(M_n > s) + \mathbb{P}_\mu(U_n > t)^K \min\left(1, \frac{1}{K! \cdot \mathbb{P}_\mu(U_n \leq t)^{K-1}}\right) \\
&\quad \leq \mathbb{P}_\mu(M_n > s) + \frac{1}{K!}\left(\frac{\mathbb{P}_\mu(U_n > t)}{\mathbb{P}_\mu(U_n \leq t)}\right)^K.
\end{aligned}$$

Second, in [3] it is not shown whether or not the variant of the Hoffmann–Jørgensen inequality (Theorem 2) can be reconciled with Theorem 1. Our result achieves this goal, thus unifying and simultaneously generalizing variants from the literature, including by Johnson and Schechtman [5], Klass and Nowicki [11] and Hitczenko and Montgomery-Smith [3].

Finally, Theorem A does not require a norm, group structure, commutativity or completeness, but is valid in the primitive mathematical setting of separable metric semigroups. Thus, the result is a significant generalization of the original inequality by Hoffmann–Jørgensen.

3. Proof of the theorem. In order to prove Theorem A, we first study basic properties of metric semigroups $\mathscr{G}$. We begin with the *triangle inequality* in $\mathscr{G}$, which is straightforward, and used without further reference:

$$d_{\mathscr{G}}(y_1y_2, z_1z_2) \leq d_{\mathscr{G}}(y_1, z_1) + d_{\mathscr{G}}(y_2, z_2), \qquad \forall y_i, z_i \in \mathscr{G}. \tag{6}$$

We also require the following lemma, which provides a work around for the "norm" in a metric semigroup, when there is no identity element.

LEMMA 5. *Given a metric semigroup $(\mathscr{G}, d_{\mathscr{G}})$, and $a, b \in \mathscr{G}$,*

$$d_{\mathscr{G}}(a, ba) = d_{\mathscr{G}}(b, b^2) = d_{\mathscr{G}}(a, ab) \tag{7}$$

is independent of $a \in \mathscr{G}$.

PROOF. Compute using the translation-invariance of $d_{\mathscr{G}}$:

$$d_{\mathscr{G}}(a, ba) = d_{\mathscr{G}}(ba, b^2a) = d_{\mathscr{G}}(b, b^2) = d_{\mathscr{G}}(ab, ab^2) = d_{\mathscr{G}}(a, ab). \qquad \square$$

Now we show the main result of the paper.

PROOF OF THEOREM A. Our proof follows in part the argument in [3]; however, we are able to streamline some of the steps and provide novel techniques that help generalize the result to its present form. For convenience, the proof is divided into steps.

Step 1. Define $K := \sum_{i=1}^k n_i$, and given $1 \leq l \leq K$, let $t'_l := t_i$ if $\sum_{j=1}^{i-1} n_j < l \leq \sum_{j=1}^{i} n_j$. Also define

$$\zeta := (2n_1 - 1)t_1 + 2\sum_{i=2}^k n_i t_i + \left(\sum_{i=1}^k n_i - 1\right)s, \qquad Y := \sum_{j=n-K+2}^{n} Y_{(j)}. \tag{8}$$

Now if $Y > (K-1)s$, then it is clear that $M_n > s$. Thus, the inequality is strengthened by replacing $\mathbb{P}_\mu(M_n > s)$ by $\mathbb{P}_\mu(Y > (K-1)s)$. (Note that this strengthening of the inequality was originally suggested in the setting of Banach spaces by Rudelson in [3].) Now set $\Omega_1 := \{\omega \in \Omega : U_n(\omega) > \zeta, Y(\omega) \leq (K-1)s\}$. Then

$$\mathbb{P}_\mu(U_n > \zeta) \leq \mathbb{P}_\mu\big(Y > (K-1)s\big) + \mathbb{P}_\mu(\Omega_1).$$

Thus, we will restrict ourselves to Ω_1. Define $m_0 = m_0(\omega) := 0$, and let $m_1(\omega) > 0$ be the smallest integer such that $d_{\mathscr{G}}(z_1, z_0 S_{m_1}(\omega)) > t_1$. Note that such an $m_1(\omega)$ exists because $t_1 \leq \zeta$ and $\omega \in \Omega_1$.

Step 2. For this step, fix $\omega \in \Omega_1$. In this step, we inductively define integers

$$m_l = m_l(\omega) \qquad \text{with } 0 = m_0 < m_1 < m_2 < \cdots < m_K \leq n$$

as follows: m_1 is as above, and given m_{l-1} for $l > 1$, define m_l to be the least integer $> m_{l-1}$ such that $d_{\mathscr{G}}(S_{m_{l-1}}, S_{m_l}) > 2t'_l$. To do so, we first *claim* that such an integer m_l exists for all $1 \le l \le K$.

To show this claim, suppose to the contrary that such an m_l does not exist (for the smallest such $l > 1$). Then for all $\beta > m_{l-1}$, $d_{\mathscr{G}}(S_{m_{l-1}}, S_\beta) \le 2t'_l$. We now make the *sub-claim* that

$$d_{\mathscr{G}}\big(z_1, z_0 S_\alpha(\omega)\big) \le t'_1 + \sum_{j=2}^{l} 2t'_j + (K-1)s \le \zeta, \qquad \forall 1 \le \alpha \le n.$$

Notice that the sub-claim contradicts the fact that we are restricted to $\omega \in \Omega_1$, thereby proving the claim. Thus, it suffices to show the sub-claim. To do so, we consider various cases: if $\alpha < m_1$, then $d_{\mathscr{G}}(z_1, z_0 S_\alpha) \le t'_1$, so we are done. Next, if $\alpha \in (m_i, m_{i+1})$ for some $0 < i < l-1$, then compute using equation (7), and that $Y \le (K-1)s$ on Ω_1:

$$\begin{aligned}
&d_{\mathscr{G}}(z_1, z_0 S_\alpha) \\
&\quad \le d_{\mathscr{G}}(z_1, z_0 S_{m_1 - 1}) + d_{\mathscr{G}}(z_0 S_{m_i - 1}, z_0 S_{m_i}) + d_{\mathscr{G}}(z_0 S_{m_i}, z_0 S_\alpha) \\
&\qquad + \sum_{j=2}^{i} \big[d_{\mathscr{G}}(z_0 S_{m_{j-1}-1}, z_0 S_{m_{j-1}}) + d_{\mathscr{G}}(z_0 S_{m_{j-1}}, z_0 S_{m_j - 1})\big] \\
&\quad \le t_1 + \sum_{j=2}^{i} \big(Y_{m_{j-1}} + 2t'_j\big) + Y_{m_i} + 2t'_{i+1} \le t'_1 + 2\sum_{j=2}^{l-1} t'_j + Y \\
&\quad \le t'_1 + 2\sum_{j=2}^{l-1} t'_j + (K-1)s.
\end{aligned}$$

There are two other cases with similar computations (hence are skipped):

- If $\alpha = m_i$ for some $i < l$, then

$$d_{\mathscr{G}}(z_1, z_0 S_\alpha) \le t'_1 + 2\sum_{j=2}^{i} t'_j + (i+1)s \le t'_1 + 2\sum_{j=2}^{l-1} t'_j + (K-1)s.$$

- If $\alpha \in (m_{l-1}, n]$, then the sub-claim follows by using that $d_{\mathscr{G}}(S_{m_{l-1}}, S_\alpha) \le 2t'_l$ from above.

Proceeding by induction on l, the above analysis in this step proves the claim about the existence of $0 = m_0(\omega) < \cdots < m_K(\omega) \le n$, for all $\omega \in \Omega_1$.

Step 3. Given a strictly increasing sequence $\mathbf{m} := (m_1, \dots, m_K)$ such that $0 = m_0 < m_1 < \cdots < m_K \le n$, define $\Omega_{\mathbf{m}}$ to be the subset of all $\omega \in \Omega_1$ such that $m_i(\omega) = m_i$ for all i. Then Ω_1 is the disjoint union of the $\Omega_{\mathbf{m}}$.

Now given $0 \leq \alpha < \beta \leq n$ and $t > 0$, define

$$
\begin{aligned}
(9) \qquad p_{\alpha,\beta,t} &:= \mathbb{P}_\mu\big(d_{\mathscr{G}}(z_0 S_\alpha, z_0 S_\beta) > 2t \geq d_{\mathscr{G}}(z_0 S_\alpha, z_0 S_j)\ \forall \alpha \leq j < \beta\big), \\
p_{\beta,t} &:= \mathbb{P}_\mu\big(d_{\mathscr{G}}(z_1, z_0 S_\beta) > t \geq d_{\mathscr{G}}(z_1, z_0 S_j)\ \forall 0 \leq j < \beta\big),
\end{aligned}
$$

where by equation (7), we may disregard the z_0's occurring in $p_{\alpha,\beta,t}$ except for $\alpha = 0$, in which case we define $z_0 S_0 := z_0$. Then by independence of the X_j [and equation (7)],

$$
\mathbb{P}_\mu(\Omega_{\mathbf{m}}) \leq p_{m_1,t_1'} \prod_{j=2}^{K} p_{m_{j-1},m_j,t_j'}.
$$

This allows us to continue the computations toward proving the result

$$
\begin{aligned}
(10) \qquad \mathbb{P}_\mu(U_n > \zeta) &\leq \mathbb{P}_\mu\big(Y > (K-1)s\big) + \mathbb{P}_\mu(\Omega_1) \\
&\leq \mathbb{P}_\mu\big(Y > (K-1)s\big) + \sum_{\mathbf{m}} p_{m_1,t_1'} \prod_{j=2}^{K} p_{m_{j-1},m_j,t_j'}.
\end{aligned}
$$

Step 4. For the next steps in the computations, we bound $\sum_{\beta=\alpha+1}^{\gamma} p_{\alpha,\beta,t}$ in two different ways, where $\alpha, \beta, \gamma \in \mathbb{N}$. First,

$$
\begin{aligned}
\sum_{\beta=\alpha+1}^{\gamma} p_{\alpha,\beta,t} &= \mathbb{P}_\mu\Big(\max_{\beta \in (\alpha,\gamma]} d_{\mathscr{G}}(S_\alpha, S_\beta) > 2t\Big) \\
&\leq \mathbb{P}_\mu\Big(\max_{\beta \in (\alpha,\gamma]} d_{\mathscr{G}}(z_1, z_0 S_\alpha) + d_{\mathscr{G}}(z_1, z_0 S_\beta) > 2t\Big) \\
&\leq \mathbb{P}_\mu(2U_\gamma > 2t) = \mathbb{P}_\mu(U_\gamma > t).
\end{aligned}
$$

(Here, U_γ is defined similar to U_n.) Similarly,

$$
\sum_{\beta=1}^{\gamma} p_{\beta,t} = \mathbb{P}_\mu\Big(\max_{\beta \in [1,\gamma]} d_{\mathscr{G}}(z_1, z_0 S_\beta) > t\Big) = \mathbb{P}_\mu(U_\gamma > t).
$$

Next, if $\mathbb{P}_\mu(U_\alpha \leq t) > 0$, then using the independence of the X_j,

$$
\begin{aligned}
&\sum_{\beta=\alpha+1}^{\gamma} p_{\alpha,\beta,t} \\
&\quad = \mathbb{P}_\mu\Big(\max_{\beta \in (\alpha,\gamma]} d_{\mathscr{G}}(S_\alpha, S_\beta) > 2t\Big) = \mathbb{P}_\mu\Big(\max_{\beta \in (\alpha,\gamma]} d_{\mathscr{G}}(S_\alpha, S_\beta) > 2t | U_\alpha \leq t\Big) \\
&\quad \leq \frac{\mathbb{P}_\mu(\max_{\alpha < \beta \leq \gamma} d_{\mathscr{G}}(z_1, z_0 S_\beta) > t \text{ and } \max_{1 \leq \beta \leq \alpha} d_{\mathscr{G}}(z_1, z_0 S_\beta) \leq t)}{\mathbb{P}_\mu(U_\alpha \leq t)} \\
&\quad = \frac{1}{\mathbb{P}_\mu(U_\alpha \leq t)} \sum_{\beta=\alpha+1}^{\gamma} p_{\beta,t}.
\end{aligned}
$$

These calculations are summarized in the following system of inequalities:

$$\begin{aligned}
&\sum_{\beta=\alpha+1}^{\gamma} p_{\alpha,\beta,t} \le \mathbb{P}_\mu(U_\gamma > t) = \sum_{\beta=1}^{\gamma} p_{\beta,t},\\
&\sum_{\beta=\alpha+1}^{\gamma} p_{\alpha,\beta,t} \le \frac{1}{\mathbb{P}_\mu(U_\alpha \le t)} \sum_{\beta=\alpha+1}^{\gamma} p_{\beta,t}.
\end{aligned}\tag{11}$$

Step 5. We now perform what is in a sense the "main step" of the computation. More precisely, we use the previous step to bound from above the following expression from equation (10):

$$\widetilde{S} := \sum_{\mathbf{m}} p_{m_1,t_1'} \prod_{j=2}^{K} p_{m_{j-1},m_j,t_j'},$$

where the summation is over all $0 < m_1 < \cdots < m_K \le n$.

For $1 \le i \le k+1$, define $s_i := \sum_{j=1}^{i-1} n_j$. Then $t_l' = t_i$ for $s_i < l \le s_i + n_i$. Suppose $k > 1$. We bound $\widetilde{S}$ via induction on k, presented here in a reverse manner. Namely, we sum first over m_j for $j \in (s_k, s_{k+1}] = (K - n_k, K]$; then over $j \in (s_{k-1}, s_k]$; and so on, reducing to the base case $k = 1$ (addressed in the next step). In the present step, we stop after one round of summation

$$\widetilde{S} = \sum_{\mathbf{m}_k} p_{m_1,t_1'} \prod_{j=2}^{s_k} p_{m_{j-1},m_j,t_j'} \cdot \sum_{\alpha_0 = m_{s_k} < \alpha_1 < \cdots < \alpha_{n_k} \le n} \prod_{j=1}^{n_k} p_{\alpha_{j-1},\alpha_j,t_k},$$

where the outer sum is over $\mathbf{m}_k := \{m_j : j \le s_k\}$. We claim that for all fixed m_j for $j \notin (s_k, s_k + n_k]$, the inner sum can be bounded above by an expression occurring in Theorem A [see (5)]. More precisely, we claim

$$\begin{aligned}
&\sum_{\alpha_0 = m_{s_k} < \alpha_1 < \cdots < \alpha_{n_k} \le n} \prod_{j=1}^{n_k} p_{\alpha_{j-1},\alpha_j,t_k}\\
&\quad \le \mathbb{P}_\mu(U_n > t_k)^{n_k} \min\left(1, \frac{1}{n_k! \cdot \mathbb{P}_\mu(U_n \le t_k)^{n_k}}\right)
\end{aligned}\tag{12}$$

(note, $k > 1$). To see why, using (11), the sum in (12) is bounded above by

$$\begin{aligned}
&\sum_{\alpha_0 = m_{s_k} < \alpha_1 < \cdots < \alpha_{n_k} \le n} \prod_{j=1}^{n_k} p_{\alpha_{j-1},\alpha_j,t_k}\\
&\quad = \sum_{\alpha_0 = m_{s_k} < \alpha_1 < \cdots < \alpha_{n_k-1} \le n} \prod_{j=1}^{n_k-1} p_{\alpha_{j-1},\alpha_j,t_k} \cdot \sum_{\alpha_{n_k} = \alpha_{n_k-1}+1}^{n} p_{\alpha_{n_k-1},\alpha_{n_k},t_k}\\
&\quad \le \sum_{\alpha_0 = m_{s_k} < \alpha_1 < \cdots < \alpha_{n_k-1} \le n} \prod_{j=1}^{n_k-1} p_{\alpha_{j-1},\alpha_j,t_k} \cdot \mathbb{P}_\mu(U_n > t_k)
\end{aligned}$$

$$\leq \mathbb{P}_\mu(U_n > t_k) \sum_{\alpha_0 = m_{s_k} < \cdots < \alpha_{n_k - 2} \leq n} \prod_{j=1}^{n_k - 2} p_{\alpha_{j-1}, \alpha_j, t_k}$$
$$\times \sum_{\alpha_{n_k - 1} = \alpha_{n_k - 2} + 1}^{n} p_{\alpha_{n_k - 2}, \alpha_{n_k - 1}, t_k}.$$

Continuing inductively, we obtain an upper bound of $\mathbb{P}_\mu(U_n > t_k)^{n_k}$.

Next, if $\mathbb{P}_\mu(U_n \leq t_k) > 0$, then we bound the sum in (12) using (11), as in the proof of [3], Theorem 1; this yields an upper bound of

$$\sum_{\alpha_0 = 0 < \alpha_1 < \cdots < \alpha_{n_k} \leq n} \prod_{j=1}^{n_k} p_{\alpha_{j-1}, \alpha_j, t_k}$$
$$\leq \frac{1}{\mathbb{P}_\mu(U_n \leq t_k)^{n_k}} \sum_{1 \leq \alpha_1 < \cdots < \alpha_{n_k} \leq n} \prod_{j=1}^{n_k} p_{\alpha_j, t_k}.$$

Since n_k distinct numbers may be arranged in $n_k!$ ways, adopting an argument in the proof of [3], Theorem 1, shows the right-hand side is at most

$$\begin{aligned} & \frac{1}{n_k!} \frac{1}{\mathbb{P}_\mu(U_n \leq t_k)^{n_k}} \left(\sum_{\beta=1}^{n} p_{\beta, t_k} \right)^{n_k} \\ & \quad = \frac{1}{n_k!} \frac{\mathbb{P}_\mu(U_n > t_k)^{n_k}}{\mathbb{P}_\mu(U_n \leq t_k)^{n_k}}. \end{aligned} \tag{13}$$

This analysis proves the claim in (12). Note as in (5), the minimum corresponds precisely to whether or not $k \in I_0$, as in the statement of the theorem. [The statement of the result also includes the case when $\mathbb{P}_\mu(U_n \leq t_k) = 0$.]

Step 6. Starting from (10), we now have a nested sum over m_j, $j \in [1, s_k]$, as the estimate obtained in (13) can be taken outside the sum over the m_j. Repeat the computation in Step 5, summing over the m_j with $j \in (s_{k-1}, s_k]$; then over $j \in (s_{k-2}, s_{k-1}]$; and so on. This yields the expression for $k = 1$:

$$\widetilde{S} \leq \prod_{1 < i \in I_0} \mathbb{P}_\mu(U_n > t_i)^{n_i} \prod_{1 < i \notin I_0} \frac{1}{n_i!} \left(\frac{\mathbb{P}_\mu(U_n > t_i)}{\mathbb{P}_\mu(U_n \leq t_i)} \right)^{n_i}$$
$$\times \sum_{\{m_j : j \in [1, n_1]\}} p_{m_1, t_1} \prod_{j=2}^{n_1} p_{m_{j-1}, m_j, t_1}.$$

It remains to find an upper bound for this last summation. To do so, follow the computations in the previous step, using equation (11). Thus, on the one hand, this summation is again at most $\mathbb{P}_\mu(U_n > t_1)^{n_1}$. On the other hand, it is bounded above, assuming that $\mathbb{P}_\mu(U_n \le t_1) > 0$, by

$$\begin{aligned}
&\frac{1}{\mathbb{P}_\mu(U_n \le t_1)^{n_1-1}} \sum_{1\le m_1<\cdots<m_{n_1}\le n} \prod_{j=1}^{n_1} p_{m_j,t_1} \\
&\quad \le \frac{1}{\mathbb{P}_\mu(U_n \le t_1)^{n_1-1}} \frac{1}{n_1!} \left(\sum_{\beta=1}^{n} p_{\beta,t_1} \right)^{n_1} \\
&\quad = \mathbb{P}_\mu(U_n > t_1)^{n_1} \cdot \frac{1}{n_1! \cdot \mathbb{P}_\mu(U_n \le t_1)^{n_1-1}},
\end{aligned}$$

and by equation (5), this completes the proof of the theorem. □

Concluding remarks. The validity of the Hoffmann–Jørgensen inequality in the metric semigroup setting suggests further work along two directions. First, the Banach space version of this inequality is an important result in the literature that is widely used in bounding sums of independent Banach space-valued random variables. Having proved Theorem A, we apply it in related work [9] to obtain similar tail bounds for sums of independent metric semigroup-valued random variables. Additionally, in [10] we study other probability inequalities for metric (semi)groups, such as the Khinchin–Kahane inequality, together with its connections to embedding abelian normed metric groups into (minimal) Banach spaces.

Acknowledgments. We thank David Montague and Doug Sparks for carefully going through an early draft of the paper and providing detailed feedback, which improved the exposition. We also thank the anonymous referee for providing technical feedback that helped clarify the proof.

REFERENCES

[1] CABELLO SÁNCHEZ, F. and CASTILLO, J. M. F. (2002). Banach space techniques underpinning a theory for nearly additive mappings. *Dissertationes Math.* (*Rozprawy Mat.*) **404** 73 pp. MR192693

[2] GRENANDER, U. (1963). *Probabilities on Algebraic Structures*. Wiley, New York. MR0206994

[3] HITCZENKO, P. and MONTGOMERY-SMITH, S. J. (2001). Measuring the magnitude of sums of independent random variables. *Ann. Probab.* **29** 447–466. MR1825159

[4] HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186. MR0356155

[5] JOHNSON, W. B. and SCHECHTMAN, G. (1989). Sums of independent random variables in rearrangement invariant function spaces. *Ann. Probab.* **17** 789–808. MR0985390

[6] KAHANE, J.-P. (1985). *Some Random Series of Functions*. *Cambridge Studies in Advanced Mathematics* **5**. Cambridge Univ. Press, London. MR0833073

[7] KHARE, A. and RAJARATNAM, B. (2014). Differential calculus on the space of countable labelled graphs. Preprint. Available at arXiv:1410.6214.

[8] KHARE, A. and RAJARATNAM, B. (2015). Integration and measures on the space of countable labelled graphs. Preprint. Available at arXiv:1506.01439.

[9] KHARE, A. and RAJARATNAM, B. (2015). Probability inequalities and tail estimates for metric semigroups. Preprint. Available at arXiv:1506.02605.

[10] KHARE, A. and RAJARATNAM, B. (2016). The Khinchin–Kahane inequality and Banach space embeddings for metric groups. Preprint. Available at arXiv:1610.03037.

[11] KLASS, M. J. and NOWICKI, K. (2000). An improvement of Hoffmann–Jørgensen's inequality. *Ann. Probab.* **28** 851–862. MR1782275

[12] KLEE, V. L. JR. (1952). Invariant metrics in groups (solution of a problem of Banach). *Proc. Amer. Math. Soc.* **3** 484–487. MR0047250

[13] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces* (*Isoperimetry and Processes*). Springer, Berlin. MR1102015

[14] LOVÁSZ, L. (2012). *Large Networks and Graph Limits*. *Colloquium Publications* **60**. Amer. Math. Soc., Providence. MR3012035

[15] RUDIN, W. (1962). *Fourier Analysis on Groups*, *Interscience*. Wiley, New York.

[16] STERNBERG, S. (1965). *Lectures on Differential Geometry*. Prentice-Hall, Englewood Cliffs. MR0268801

DEPARTMENT OF MATHEMATICS
INDIAN INSITUTE OF SCIENCE
BANGALORE 560012
INDIA
E-MAIL: khare@iisc.ac.in

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, DAVIS
ONE SHIELDS AVENUE
DAVIS, CALIFORNIA 95616
USA
E-MAIL: brajaratnam01@gmail.com