Small and Pseudo-Small Sets for Markov Chains

by

Gareth O. Roberts* and Jeffrey S. Rosenthal**

(April 2000; last revised December 2000.)

In this paper we examine the relationship between small sets and their generalisation, pseudo-small sets. We consider conditions which imply the equivalence of the two notions, and give examples where they are definitely different. We give further examples where sets are both pseudo-small and small, but the minorisation constants implied by the two notions are different. Applications of recent computable bounds results are given and extended. We also give a result linking the ideas of monotonicity and minorisation. Specifically we demonstrate that if a non-monotone chain satisfies a minorisation condition, and furthermore is stochastically dominated by a monotone chain which satisfies a Lyapunov drift condition, then a probability construction exists which incorporates both the bounding process and the minorisation condition.

Keywords. Coupling, convergence rates, small set, minorisation condition, total variation distance.

1. Introduction.

A well-known concept in the theoretical study of Markov chains on general state spaces is the property of *small set*, or *minorisation condition*. This is the property that all of the transition probability distributions $P^{n_0}(x, \cdot)$, for some fixed $n_0 \in \mathbf{N}$ and for all x in some subset C, all have a certain non-zero component in common. In symbols, $P^{n_0}(x, \cdot) \geq \epsilon\, \nu(\cdot)$ for all $x \in C$. Such small sets arise often in the theoretical study of Markov chain Monte Carlo (MCMC) algorithms; see e.g. Smith and Roberts (1993), Tierney (1994), Gilks, Richardson and Spiegelhalter (1996), and Roberts and Rosenthal (1998).

* Department of Mathematics and Statistics, Fylde College, Lancaster University, Lancaster, LA1 4YF, England. Internet: `g.o.roberts@lancaster.ac.uk`.

** Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S 3G3. Internet: `jeff@math.toronto.edu`. Supported in part by NSERC of Canada.

If the small set C is hit infinitely often with probability 1, then it allows for the construction of *regeneration times* (cf. Athreya and Ney, 1978; Nummelin, 1978, 1984; Asmussen, 1987; Mykland, Tierney, and Yu, 1995). If the mean interarrival time is finite, then such regeneration times guarantee the existence of a *stationary distribution*, and furthermore provide some control (in terms of hitting times of C) over how quickly the chain converges to this stationary distribution.

Small sets also allow for the construction of *couplings* (cf. Lindvall, 1992; Meyn and Tweedie, 1993; Rosenthal, 1995a; see Appendix for details), whereby two different copies of the chain become equal with positive probability (since they may both update from the same distribution $\nu(\cdot)$). Such couplings also provide bounds on convergence rates to stationary distributions (though without guaranteeing the *existence* of a stationary distribution), through the *coupling inequality*.

Finally, since small sets provide a condition which holds for *all* elements of C simultaneously, they can also be used to construct *coalescence* (see e.g. Murdoch and Green, 1998), whereby copies of the chain started at *all* elements of the state space *all* become equal simultaneously. This is especially important in exact sampling schemes such as *coupling from the past* (Propp and Wilson, 1996) and *Fill's algorithm* (Fill, 1998; Fill, Machida, Murdoch, and Rosenthal, 1999).

In a different direction, small sets can be used to construct *shift-couplings* (cf. Aldous and Thorrison, 1993; Roberts and Rosenthal, 1997), whereby two copies of the chain become equal at two *different* times. This can be used to provide bounds on the convergence rate to stationarity of *ergodic average* distributions.

A notion related to but weaker than small set is that of *pseudo-small set*, or *pseudo-minorisation condition*. This notion was introduced formally in Roberts and Rosenthal (1996), though the idea underlying it may have been observed earlier. The idea here is that every pair of points $(x, y) \in C \times C$ has a component in common, but that common component may vary depending on the pair chosen. In symbols, $P^{n_0}(x, \cdot) \geq \epsilon\, \nu_{xy}(\cdot)$ and $P^{n_0}(y, \cdot) \geq \epsilon\, \nu_{xy}(\cdot)$ for all pairs (x, y), though here ν_{xy} depends on the choice of x and y.

The pseudo-minorisation condition does *not* immediately provide notions such as regeneration, the existence of a stationary distribution, coalescence, or a shift-coupling construction. However, the pseudo-minorisation condition is perfectly adequate for ordinary (pairwise) coupling constructions, which always consider just two chains at a time (see Appendix for details). That simple observation provides the basis for the current paper.

This paper is organised as follows. In Section 3, we develop analogues of previous convergence-rate results in terms of pseudo-small sets. In Section 4, we provide specialised versions of these results for countable and absolutely-continuous chains. Section 5 presents a number of different examples of small and pseudo-small sets. In Section 6, we consider the relationship between small and pseudo-small sets, and prove that for ϕ-irreducible, aperiodic Markov chains with countably-generated σ-algebras, all pseudo-small subsets are in fact small (though perhaps with much worse values of n_0 and ϵ). Section 7 considers what can go wrong when assumptions of ϕ-irreducibility and aperiodicity are relaxed. Finally, Section 8 presents a result which applies convergence results for stochastically monotone chains to non-monotone chains which are instead *bounded* by monotone chains.

The paper closes with an Appendix which reviews the traditional pairwise coupling construction based on small sets, and describes how the construction can be modified to be used for pseudo-small sets.

2. Definitions.

Let $\{X_n\}$ be a Markov chain on a state space $\mathcal{X}$, having transition probabilities $P(x, \cdot)$. We begin with a standard definition.

Definition. A set $C \subseteq \mathcal{X}$ is *small* (or, (n_0, ϵ, ν)-small) if there is $n_0 \in \mathbf{N}$, $\epsilon > 0$, and a probability measure ν, such that

$$P^{n_0}(x, \cdot) \;\geq\; \epsilon\nu(\cdot)\,, \qquad x \in C\,. \tag{1}$$

The existence of small sets for ϕ-irreducible Markov chains is proved in Jain and Jameson (1967) and Orey (1971); see Meyn and Tweedie (1993) for a modern exposition. (Recall that a Markov chain is *ϕ-irreducible* if there is a non-zero measure ϕ on $\mathcal{X}$, such that for any subset A with $\phi(A) > 0$, there is positive probability of hitting A starting from any $x \in \mathcal{X}$. See e.g. Meyn and Tweedie, 1993, for this and other basic Markov chain definitions.)

Small sets have many uses. The one most relevant to the current paper is for *pairwise coupling constructions*. Briefly, we can construct two copies of the Markov chain (one started in the arbitrary initial distribution, the other started in the stationary distribution) such that, each time they are both in the small set C, they have probability ϵ of coupling n_0 iterations later. For formal details, see the Appendix.

We now introduce a related, but weaker, notion than that of a small set.

Definition. A set $C \subseteq \mathcal{X}$ is *pseudo-small* (or, $(n_0, \epsilon, \{\nu_{xy}\})$-pseudo-small) if there is $n_0 \in \mathbf{N}$ and $\epsilon > 0$ such that for all $(x, y) \in C \times C$, there is a probability measure ν_{xy} with

$$P^{n_0}(x, \cdot) \wedge P^{n_0}(y, \cdot) \geq \epsilon \nu_{xy}(\cdot) \,. \tag{2}$$

(Note that ((???)) is shorthand for the two equations $P^{n_0}(x, A) \geq \epsilon \nu_{xy}(A)$ and $P^{n_0}(y, A) \geq \epsilon \nu_{xy}(A)$ for all measurable sets A.)

Obviously, any small set is also pseudo-small, with the same n_0 and ϵ, and with $\nu_{xy} = \nu$ for all pairs (x, y). The primary motivation for pseudo-small sets is that the usual pairwise coupling construction for small sets can be used essentially without change for pseudo-small sets (see Appendix). This means that any result proved using the small set pairwise coupling construction has an immediate analogues for pseudo-small sets, as we now explore.

3. General pseudo-small convergence results.

As described above, for any convergence bounds involving the ordinary (pairwise) coupling construction as in the Appendix, the bound remains true if a minorisation condition is replaced by a corresponding pseudo-minorisation condition. Thus, any coupling-based convergence result which uses a small set can immediately be “transformed” into a corresponding result involving pseudo-small sets.

In particular, because of the purely coupling proof of the following result for small sets (cf. Rosenthal, 1993; Meyn and Tweedie, 1993, Theorem 16.2.4), we have

Proposition 1. *Let $P(x, \cdot)$ be the transition probabilities for a Markov chain on a state space $\mathcal{X}$, with stationary distribution $\pi(\cdot)$. If the entire state space $\mathcal{X}$ is (n_0, ϵ)-pseudo-small, then*

$$\|P^n(x, \cdot) - \pi(\cdot)\| \ \leq \ (1-\epsilon)^{\lfloor n/n_0 \rfloor}, \qquad n \in \mathbf{N},$$

independent of the initial value $x \in \mathcal{X}$.

Here and throughout, $\lfloor r \rfloor$ is the greatest integer not exceeding the real number r, and $\|P^n(x, \cdot) - \pi(\cdot)\| \equiv \sup_{A \subseteq \mathcal{X}} |P^n(x, A) - \pi(A)|$ represents the *total variation distance* between the actual distribution of the Markov chain (after n steps, when initially started at the point x), and the stationary distribution $\pi(\cdot)$.

Similarly, by transforming Proposition 1 of Cowles and Rosenthal (1998; based on Theorem 12 of Rosenthal, 1995a; see also Roberts and Tweedie, 1999), we obtain

Proposition 2. *Let $P(x, \cdot)$ be the transition probabilities for a Markov chain on a state space $\mathcal{X}$, with stationary distribution $\pi(\cdot)$. Suppose for some function $V : \mathcal{X} \to [0, \infty)$, some $\lambda < 1$ and $\Lambda < \infty$, some $\epsilon > 0$, some positive integers m and k_0, and some $d > \frac{2\Lambda}{1-\lambda}$, we have the drift condition*

$$\mathbf{E}(V(X_m) \mid X_0 = x) \leq \lambda V(x) + \Lambda, \qquad x \in \mathcal{X},$$

and also that the set $\{x \in \mathcal{X}; V(x) \leq d\}$ is (mk_0, ϵ)-pseudo-small. Then for any $0 < r < 1$ and $M > 0$, we have

$$\|\mathcal{L}(X_k) \ - \ \pi\| \ \leq \ (1-\epsilon)^{\lfloor rk/mk_0 \rfloor} \ + \ C_0 \, (\alpha A)^{-1} \left(\alpha^{-(1-rk_0)} A^r\right)^{\lfloor k/m \rfloor}, \qquad k \in \mathbf{N},$$

where

$$\alpha^{-1} \ = \ \frac{1 + 2M\Lambda + M\lambda d}{1 + Md}; \quad A \ = \ 1 + 2(\lambda M d + M\Lambda); \quad C_0 \ = \ \left(1 + \frac{M\Lambda}{1-\lambda} + M\mathbf{E}(V(X_0))\right).$$

If furthermore it is known that $V(x) \geq 1$ for all $x \in \mathcal{X}$, then it suffices that $d > \frac{2\Lambda}{1-\lambda} - 1$, and these values may be improved slightly to

$$\alpha^{-1} \ = \ \lambda + \frac{M\Lambda + (1-\lambda)(1-M)}{1 + \frac{M}{2}(d-1)}; \qquad A \ = \ M(\lambda d + \Lambda) + (1-M);$$

$$C_0 \;=\; \frac{M}{2}\left(\frac{\Lambda}{1-\lambda}+\mathbf{E}(V(X_0))\right)+(1-M)\,.$$

For example, taking $M=1$ in the $V(x)\geq 1$ case, we obtain the simplification

$$\alpha^{-1} \;=\; \lambda+\frac{2\Lambda}{d+1}\,; \quad A \;=\; \lambda d+\Lambda\,; \quad C_0 \;=\; \frac{1}{2}\left(\frac{\Lambda}{1-\lambda}+\mathbf{E}(V(X_0))\right)\,.$$

Recall now that a Markov chain is *stochastically monotone* with respect to an ordering $\preceq$ on $\mathcal{X}$ if, for all fixed $\mathbf{z}$, we have that $\mathbf{P}(\mathbf{X}_1 \preceq \mathbf{z}|\mathbf{X}_0=\mathbf{x}_1) \geq \mathbf{P}(\mathbf{X}_1 \preceq \mathbf{z}|\mathbf{X}_0=\mathbf{x}_2)$ whenever $\mathbf{x}_1 \preceq \mathbf{x}_2$. For such chains, if $X_0 \succeq Y_0$, then it is possible (cf. Kamae et al., 1977; Lindvall, 1992, p. 134) to simultaneously construct $\{X_k\}$ and $\{Y_k\}$ so that $X_k \succeq Y_k$ for all k. Intuitively, for stochastically monotone chains and small sets of the form $C=\{x\in\mathcal{X};\ x\preceq c\}$, it is easier to prove convergence bounds, since if $X_k \succeq Y_k$ for all k, then $(X_k,Y_k)\in C\times C$ whenever $X_k\in C$.

Therefore, transforming Theorem 2.2 of Roberts and Tweedie (2000; which builds on the work of Lund and Tweedie, 1996 and Lund, Meyn, and Tweedie, 1996), we obtain the following. We write $\nu(V)$ for the expected value of V with respect to ν, and write $\mathbf{E}_x^{\pi}(V)$ for the expected value of V with respect to the *stochastic majorant* (with respect to $\preceq$) of a point mass at x and the stationary distribution $\pi(\cdot)$, defined by

$$\mathbf{E}_x^{\pi}(\mathbf{1}_{(-\infty,y]}) \;=\; \min\left[\mathbf{1}_{(-\infty,y]}(x),\ \pi\left((-\infty,y]\right)\right]\,.$$

Proposition 3. *Let $P(x,\cdot)$ be the transition probabilities for a stochastically monotone Markov chain on a totally ordered state space $\mathcal{X}$, with stationary distribution $\pi(\cdot)$. Let $C\subseteq\mathcal{X}$ be $(1,\epsilon)$-pseudo-small, where $C=\{x\in\mathcal{X};\ x\preceq c\}$ for some fixed $c\in\mathcal{X}$, and let $V:\mathcal{X}\to[1,\infty)$ be such that*

$$\mathbf{E}(V(X_1)\mid X_0=x) \;\leq\; \lambda V(x)+b\mathbf{1}_C(x)\,, \qquad x\in\mathcal{X}$$

for some $\lambda<1$ and $0\leq b<\infty$. Then for $n>\log\mathbf{E}_x^{\pi}(V)\,/\,\log(\lambda^{-1})$, we have

$$\|P^n(x,\cdot)-\pi(\cdot)\| \;\leq\; K(n+\eta-\xi)\rho^n\,, \qquad n\in\mathbf{N}\,.$$

Here

$$K = \frac{e\epsilon(1-\epsilon)^{-\xi/\eta}}{\eta},$$

$$\xi = \frac{\log \mathbf{E}_x^\pi(V)}{\log(\lambda^{-1})}, \qquad \eta = \frac{\log\left(\frac{\lambda s + b - \epsilon}{\lambda(1-\epsilon)}\right)}{\log(\lambda^{-1})},$$

$s = \sup\{V(z);\, z \preceq x\}$, *and* $\rho = (1-\epsilon)^{\eta^{-1}}$.

In summary, any total variation distance convergence result based on pairwise coupling, which makes use of small sets, can immediately be transformed into a corresponding result using the weaker condition of pseudo-small sets.

Remark. Of course, there are many Markov chains which do not converge at all in total variation distance, so that neither small nor pseudo-small sets can be used in the above manner. Rather, other distance measures and other convergence techniques must be used; see e.g. Su (1998).

4. Countable pseudo-small state spaces.

For Markov chains on countable state spaces, certain more explicit formulae are available. We begin with the standard

Proposition 4. *Consider a Markov chain on a countable state space $\mathcal{X}$, and let $C \subseteq \mathcal{X}$ be any subset. Then for any $n_0 \in \mathbf{N}$, the subset C is (n_0, ϵ_{n_0})-small with*

$$\epsilon_{n_0} \;=\; \sum_{y \in \mathcal{X}} \inf_{x \in C} P^{n_0}(x, \{y\}).$$

(Of course, we may have $\epsilon_{n_0} = 0$ for all $n_0 \in \mathbf{N}$.) Furthermore, C is not (n_0, ϵ')-small for any $\epsilon' > \epsilon_{n_0}$.

Proof. Define the probability measure $\nu(\cdot)$ by

$$\nu(\{z\}) \;=\; \frac{\inf_{x \in C} P^{n_0}(x, \{z\})}{\sum_{y \in \mathcal{X}} \inf_{x \in C} P^{n_0}(x, \{y\})}.$$

Then it is verified that $P^{n_0}(x, \cdot) \geq \epsilon_{n_0}\, \nu(\cdot)$ for all $x \in C$, with ϵ_{n_0} as above.

Furthermore, if there were some $\epsilon' > \epsilon_{n_0}$ and some other probability measure $\nu'(\cdot)$ such that $P^{n_0}(x,\cdot) \geq \epsilon'\, \nu'(\cdot)$ for all $x \in C$, then we could find $z \in \mathcal{X}$ with $\epsilon'\, \nu'(\{z\}) > \epsilon_{n_0}\, \nu(\{z\})$. But $\epsilon_{n_0}\, \nu(\{z\}) = \inf_{x\in C} P^{n_0}(x,\{z\})$, so this gives a contradiction. ∎

Using the concept of pseudo-small sets, we can improve the above result to

Proposition 5. *Consider a Markov chain on a countable state space $\mathcal{X}$, and let $C \subseteq \mathcal{X}$ be any subset. Then for any $n_0 \in \mathbf{N}$, the subset C is (n_0, ϵ_{n_0})-pseudo-small with*

$$\epsilon_{n_0} \;=\; \inf_{x,y\in C} \sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})]\,.$$

(Of course, we may have $\epsilon_{n_0} = 0$ for all $n_0 \in \mathbf{N}$.) Furthermore, C is not (n_0, ϵ')-pseudo-small for any $\epsilon' > \epsilon_{n_0}$.

Proof. For $x, y \in C$, define the probability measure $\nu_{xy}(\cdot)$ by

$$\nu_{xy}(\{w\}) \;=\; \frac{\min[P^{n_0}(x,\{w\}), P^{n_0}(y,\{w\})]}{\sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})]}\,.$$

Then it is verified that $P^{n_0}(x,\cdot) \geq \epsilon_{n_0}\, \nu_{xy}(\cdot)$ and $P^{n_0}(y,\cdot) \geq \epsilon_{n_0}\, \nu_{xy}(\cdot)$ for all $x, y \in C$, with ϵ_{n_0} as above.

Furthermore, if there were some $\epsilon' > \epsilon_{n_0}$ and some other probability measures $\nu'_{xy}(\cdot)$ such that $P^{n_0}(x,\cdot) \geq \epsilon'\, \nu'_{xy}(\cdot)$ and $P^{n_0}(y,\cdot) \geq \epsilon'\, \nu'_{xy}(\cdot)$ for all $x, y \in C$, then we could find $x, y \in C$ with $\epsilon' > \sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})]$. We could then find $w \in \mathcal{X}$ with $\nu'_{xy}(\{w\}) \geq \nu_{xy}(\{w\})$. It follows that

$$\epsilon'\, \nu'_{xy}(\{w\}) \geq \epsilon'\, \nu_{xy}(\{w\}) > \sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})]\, \nu_{xy}(\{w\})$$

$$= \min[P^{n_0}(x,\{w\}), P^{n_0}(y,\{w\})]\,,$$

giving a contradiction. ∎

Combining the above proposition with Proposition (???), we obtain the following corollary (which also follows from Dobrushin, 1956, pp. 71 and 334).

Corollary 6. *Consider a Markov chain on a countable state space $\mathcal{X}$, with stationary distribution $\pi(\cdot)$. Then for any $n_0 \in \mathbf{N}$, and any $x \in \mathcal{X}$, we have*

$$\|P^n(x,\cdot) - \pi(\cdot)\| \;\leq\; (1-\epsilon_{n_0})^{\lfloor n/n_0 \rfloor}\,, \qquad n \in \mathbf{N}\,,$$

where

$$\epsilon_{n_0} \;=\; \inf_{x,y\in\mathcal{X}} \sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})]\,.$$

(Again, we may have $\epsilon_{n_0} = 0$ for all $n_0 \in \mathbf{N}$.)

As a special case of Corollary (???), we obtain an alternate proof of a special case of a result of P. Bickel.

Corollary 7. *(Bickel, 1999) Consider a Markov chain on a finite state space $\mathcal{X}$, with $|\mathcal{X}| = k$. For $n_0 \in \mathbf{N}$ and $x \in \mathcal{X}$, let $h_{n_0}(x) = \#\{y \in \mathcal{X}\,;\; P^{n_0}(x,\{y\}) > 0\}$. Suppose that for some $n_0 \in \mathbf{N}$, we have $h_{n_0}(x) > \frac{k}{2}$ for all $x \in \mathcal{X}$. Then the Markov chain is uniformly ergodic.*

Proof. Since $h_{n_0}(x) > \frac{k}{2}$ for all $x \in \mathcal{X}$, it follows easily that

$$\sum_{z\in\mathcal{X}} \min[P^{n_0}(x,\{z\}), P^{n_0}(y,\{z\})] > 0 \qquad \text{for all } x,y \in \mathcal{X}\,.$$

Hence, with ϵ_{n_0} as in the previous corollary, we have $\epsilon_{n_0} > 0$. The result now follows from the previous corollary. ■

Remark. Of course, Corollary 7 can also be proved by showing that the chain is irreducible and aperiodic, and then using standard theory. In fact, Bickel (1999) proves more than Corollary 7, showing that the chain's convergence rate can be controlled by $\max_j \sum_{i=1}^{|\mathcal{X}|} |P^{n_0}(i,\{j\}) - \text{median}_k P^{n_0}(k,\{j\})|$.

By analogy to Corollary (???), we obtain a similar result for continuous spaces when the transition distributions all have a density. (We omit the proof.)

Proposition 8. *Consider a Markov chain on a general state space $\mathcal{X}$, with stationary distribution $\pi(\cdot)$. Suppose that for some $n_0 \in \mathbf{N}$, we have that $P^{n_0}(x, dy) = f_{n_0,x}(y)m(dy)$ for all $x \in \mathcal{X}$, for some fixed σ-finite measure $m(\cdot)$ and some density functions $f_{n_0,x}$. Then*

$$\|P^n(x,\cdot) - \pi(\cdot)\| \ \leq \ (1 - \epsilon_{n_0})^{\lfloor n/n_0 \rfloor}\,, \qquad n \in \mathbf{N}\,,$$

where

$$\epsilon_{n_0} \ = \ \inf_{x,y \in \mathcal{X}} \int \min[f_{n_0,x}(z),\ f_{n_0,y}(z)]\ m(dz)\,.$$

5. Examples.

It is sometimes straightforward to construct pseudo-small sets, but far more difficult to directly construct small sets. Alternatively, sometimes it is easy to construct both, but the pseudo-small construction gives much better convergence bounds. This section presents a number of different examples to illustrate this.

Example #1: A simple illustration.

For a simple example, consider the (non-reversible) Markov chain on $\mathcal{X} = \{1,2,3\}$ with stationary distribution the uniform distribution $U(\cdot)$ on $\mathcal{X}$, and with transition probabilities given by

$$P \ = \ \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}\,.$$

Then the entire space $\mathcal{X}$ is $(1, \frac{1}{2})$-pseudo-small (with $\nu_{xy} = \delta_{z(x,y)}$ for appropriate points $z(x,y)$). Hence, pseudo-smallness gives a convergence bound of

$$\|P^n(x,\cdot) - U(\cdot)\| \ \leq \ 0.5^n\,, \qquad n \in \mathbf{N}\,,$$

On the other hand, $\mathcal{X}$ is *not* $(1,\epsilon)$-small for any $\epsilon > 0$. Instead, the chain is only $(2, \frac{3}{4})$-small (with $\nu = U$). Hence, smallness gives a convergence bound of

$$\|P^n(x,\cdot) - U(\cdot)\| \ \leq \ 0.25^{\lfloor n/2 \rfloor}\,, \qquad n \in \mathbf{N}\,.$$