How old is the potential method in coding theory?

João Ribeiro

October 29, 2026

Guruswami, Håstad, Sudan, and Zuckerman [GHSZ02] proved that there exist binary linear codes of rate $1-h(\rho)-\varepsilon$ that are $(\rho, L = O(1/\varepsilon))$-list-decodable, matching the known performance of uniformly random codes. This is a notable result in coding theory, and part of what makes it stand out is the beautiful underlying proof strategy.

At a high level, the argument first sets up a *potential* $\Phi_{\mathcal{C}}$ of a linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$ that captures the list-decoding capabilities of $\mathcal{C}$, in the sense that a low-potential code has small list size. Then, it constructs a random linear code of dimension $k = (1-h(\rho)-\varepsilon)n$ step-by-step, starting with $\mathcal{C}_0 = \{0\}$ and adding a uniformly random vector v_i to the basis in the i-th step, and tracks how the potential grows in expectation in each step.

A natural first choice for the potential would be

$$\Phi_{\mathcal{C}} = \max_{y \in \mathbb{F}_2^n} |B_{\rho n}(y) \cap \mathcal{C}|,$$

where $B_{\rho n}(y)$ denotes the radius-ρn Hamming ball in $\mathbb{F}_2^n$ centered at y. However, studying the expected growth of this potential is complicated. Instead, one considers the potential

$$\Phi_{\mathcal{C}} = 2^{-n} \sum_{y \in \mathbb{F}_2^n} 2^{\varepsilon n |B_{\rho n}(y) \cap \mathcal{C}|}.$$

This potential still heavily penalizes y's with not-small list size, and turns out to be easier to analyze. At a high level, this is because we can write $\Phi_{\mathcal{C}} = \mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}}(y)]$ for $f_{\mathcal{C}}(y) = 2^{\varepsilon n |B_{\rho n}(y) \cap \mathcal{C}|}$, where $y \sim \mathbb{F}_2^n$ means that y is sampled uniformly at random from $\mathbb{F}_2^n$, and, furthermore, for any $v \notin \mathcal{C}$ we have

$$f_{\mathcal{C}+\mathrm{span}(v)}(y) = f_{\mathcal{C}}(y) \cdot f_{\mathcal{C}}(y+v).$$

Combining these observations with linearity of expectation, it is not difficult to show that, for any fixed linear code $\mathcal{C}_i$ of dimension i and $\mathcal{C}_{i+1}$ obtained by adding to the basis $\{v_1, \ldots, v_i\}$ of $\mathcal{C}_i$ a vector v_{i+1} uniformly at random from $\mathbb{F}_2^n \setminus \mathrm{span}(v_1, \ldots, v_i)$, we have

$$\mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n \setminus \mathrm{span}(v_1, \ldots, v_i)}[\Phi_{\mathcal{C}_{i+1}}] \leq \frac{\Phi_{\mathcal{C}_i}^2}{1 - 2^{i-n}}.$$

This means that in each step there is *some* choice of a new linearly independent basis vector v_{i+1} guaranteeing that

$$\Phi_{\mathcal{C}_{i+1}} \leq \frac{\Phi_{\mathcal{C}_i}^2}{1 - 2^{i-n}}.$$

In turn, this means that

$$\Phi_{\mathcal{C}_k} \leq \frac{\Phi_{\mathcal{C}_0}^{2^k}}{\prod_{i=0}^{k-1} (1 - 2^{i-n})^{2^{k-1-i}}} \leq 2 \cdot \Phi_{\mathcal{C}_0}^{2^k}.$$

This high-level potential-based approach, of setting up an appropriate potential and studying how it evolves as we add random basis vectors one-by-one, has proved quite fruitful in coding theory. So, I think that it is fair to call this a *method*.

Was GHSZ the first application of the potential method in coding theory? If not, what was the first instance? The discussion in GHSZ does not offer any clues. It turns out that the potential method has been present in coding theory since at least *1962*, nearly 40 years before GHSZ, although disguised under different language and applied to a different, but related, problem.

How did I stumble upon this? With my collaborators, we spent a considerable amount of time thinking about list-decoding of random linear codes at rates *above capacity* and other related problems like list-recovery [DLM+26]. In this regime the list size can never be small, and the natural goal for list-decoding above capacity is instead to ensure that every radius-ρn Hamming ball contains the "expected" number of codewords. More precisely, if we aim for rate $R > 1 - h(\rho)$, then we want to show that a random binary linear code $\mathcal{C}$ of rate R satisfies

$$|B_{\rho n}(y) \cap \mathcal{C}| \approx \frac{|B_{\rho n}(0)|}{2^{(1-R)n}}$$

for all $y \in \mathbb{F}_2^n$, meaning that it matches the behavior of a uniformly random (unstructured) code of the same rate. This is a stronger guarantee than $\mathcal{C}$ being *ρ-covering*, which corresponds to showing that

$$|B_{\rho n}(y) \cap \mathcal{C}| \geq 1 \tag{1}$$

for all $y \in \mathbb{F}_2^n$.

There is a large body of work on covering codes. While thinking about list-decoding above capacity, I was leafing through the gigantic covering codes book of Cohen, Honkala, Litsyn, and Lobstein [CHLL97], and bumped into their Theorem 12.3.2. Essentially, this theorem claims a more precise version of the fact that for every $\varepsilon > 0$ there exists a binary linear code of rate $1 - h(\rho) + \varepsilon$ that is ρ-covering.

Here is a reinterpretation of the argument in the CHLL book, in GHSZ-style language. For a linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$, consider the potential

$$\Phi_{\mathcal{C}} = 2^{-n}|\{y \in \mathbb{F}_2^n : B_{\rho n}(y) \cap \mathcal{C} = \varnothing\}| = \Pr_{y \sim \mathbb{F}_2^n}[y \text{ is not covered by } \mathcal{C}].$$

Clearly, if $\Phi_{\mathcal{C}} < 2^{-n}$ then $\mathcal{C}$ must be ρ-covering. Note that, similarly to GHSZ, we may write $\Phi_{\mathcal{C}} = \mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}}(y)]$ for $f_{\mathcal{C}}(y) = \mathbf{1}_{\{B_{\rho n}(y) \cap \mathcal{C} = \varnothing\}}$, and

$$f_{\mathcal{C}+\mathrm{span}(v)}(y) = f_{\mathcal{C}}(y) \cdot f_{\mathcal{C}}(y+v).$$

These properties make the argument go through as in GHSZ, as we now show. Take $\mathcal{C}_0 = \{0\}$ and consider $\mathcal{C}_i$ obtained by adding a uniformly random vector v_i to the basis of $\mathcal{C}_{i-1}$. We have

$$\Phi_{\mathcal{C}_0} = 2^{-n}(2^n - |B_{\rho n}(0)|) \leq 1 - 2^{-n(1-h(\rho)-o(1))}.$$

Furthermore, for any fixed $\mathcal{C}_i$ we have

$$\begin{aligned}
\mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n}[\Phi_{\mathcal{C}_{i+1}}] &= \mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n}[\mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i + \mathrm{span}(v_{i+1}}(y)]] \\
&= \mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n}[\mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i}(y) \cdot f_{\mathcal{C}_i}(y + v_{i+1})]] \\
&= \mathbb{E}_{y \sim \mathbb{F}_2^n}[\mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i}(y) \cdot f_{\mathcal{C}_i}(y + v_{i+1})]] \\
&= \mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i}(y) \cdot \mathbb{E}_{v_{i+1} \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i}(y + v_{i+1})]]
\end{aligned}$$

$$= \mathbb{E}_{y \sim \mathbb{F}_2^n}[f_{\mathcal{C}_i}(y) \cdot \Phi_{\mathcal{C}_i}]$$
$$= \Phi_{\mathcal{C}_i}^2.$$

Therefore, there always exists a choice of v_{i+1} such that $\Phi_{\mathcal{C}_{i+1}} \leq \Phi_{\mathcal{C}_i}^2$ (note that here we are not worried about v_{i+1} landing in $\text{span}(v_1, \ldots, v_i)$, since a smaller code is better for us). Taking these choices over $k = (1 - h(\rho) + \varepsilon)n$ steps, we end up with a binary linear code $\mathcal{C}_k$ of rate at most $1 - h(\rho) + \varepsilon$ such that

$$\Phi_{\mathcal{C}_k} \leq \Phi_{\mathcal{C}_0}^{2^k} \leq (1 - 2^{-n(1-h(\rho)-o(1))})^{2^k} \leq e^{-2^{-n(1-h(\rho)-o(1))} \cdot 2^k} \leq e^{-2^{n(\varepsilon - o(1))}}.$$

For large enough n, we have $e^{-2^{n(\varepsilon-o(1))}} < 2^{-n}$, and so $\mathcal{C} = \mathcal{C}_k$ is ρ-covering.

The CHLL book has a nice (but, as we shall see, incomplete) historical note on developments surrounding this theorem [CHLL97, Section 12.9]. It attributes Theorem 12.3.2 to a 1983 paper of Cohen [Coh83], "with a similar result appearing in an unpublished thesis by Goblick," focusing on rate distortion theory. There is no pointer to a specific theorem in Goblick's 1962 PhD thesis [Gob62], but it seems that Theorem 4.2 is the "similar result" in question. Its statement and proof, suitably reinterpreted, essentially correspond to the potential method approach to covering codes laid out above. Just have a look at Figures 1 and 2, which present parts of Goblick's proof. Even without full context, we see the potential method in action.

We may then write Q_{j+1}, the probability of the source producing a point which cannot be encoded with r or fewer errors with the new code of j+1 generators,

$$Q_{j+1} \leq 1 - \frac{N_{j+1}}{2^n} = 1 - 2\frac{N_j}{2^n} + \left(\frac{N_j}{2^n}\right)^2 = \left(1 - \frac{N_j}{2^n}\right)^2$$

$$= Q_j^2$$

Figure 1: Part of the proof of Theorem 4.2 in Goblick's PhD thesis. Here, Q_j corresponds to $\Phi_{\mathcal{C}_j}$.

As far as I could tell, partially based on the (incomplete) historical note in the CHLL book [CHLL97, Section 12.9], the potential method developed as follows:

- 1962: Goblick [Gob62, Theorem 4.2] uses the potential method for the first time to show that there exists a binary linear code of rate $\approx 1 - h(\rho)$ that is ρ-covering, with potential $\Phi_{\mathcal{C}} = \Pr_{y \sim \mathbb{F}_2^n}[y \text{ not covered by } \mathcal{C}]$. This result and its proof were later included in Berger's 1971 book on rate distortion theory [Ber71, Theorem 6.2.2].
- 1983: Cohen [Coh83] independently obtains a similar result with the same approach.
- 1985: Cohen and Frankl [CF85] extend the potential method of Goblick [Gob62] and Cohen [Coh83] to non-binary fields. Like the GHSZ result for list-decoding, the original argument [Gob62, Coh83] only worked over $\mathbb{F}_2$. Remarkably, applications of the potential method to list-decoding remain restricted to $\mathbb{F}_2$. The reason may be the fact that covering is a "monotone property", in the sense that once y is covered by $\mathcal{C}$, then adding more vectors to $\mathcal{C}$ will not "uncover" y later. This simplifies the analysis of the step-by-step code construction.

Since we have defined in Eq. 4.12

$$Q_0 = 1 - \sum_{i=0}^{r} \binom{n}{i} 2^{-n}$$

and now the recursion relation $Q_{j+1} = Q_j^2$, our inductive proof is complete and we may write Q_k for a group code with k generators as

$$Q_k = [\, 1 - 2^{-n} \sum_{i=0}^{r} \binom{n}{i}]^{2^k},$$

where Q_k is the probability that the source produces a point which cannot be encoded with r or fewer errors.

Figure 2: Another part of the proof of Theorem 4.2 in Goblick's PhD thesis.

- 1986: Delsarte and Piret [DP86] give a simplified proof of Goblick's result and note that the problem of whether most binary linear codes of rate $1 - h(\rho) + \varepsilon$ are ρ-covering remains open.
- 1987: Blinovsky [Bli87] shows that with high probability a random linear code of rate $1 - h(\rho) + \varepsilon$ is ρ-covering using another variant of the potential method (in fact, he shows this over non-binary fields too). Levitin [Lev88] independently announced the same result in a "recent results" session at ISIT 1988 (this is not mentioned in [CHLL97], but has been mentioned elsewhere [CG90, Section II]).

  The CHLL book has a nice presentation of Blinovsky's proof [CHLL97, Theorem 12.3.5]. Roughly speaking, it first uses pairwise independence and Chebyshev's inequality to argue that a random linear code $\mathcal{C}'$ of dimension $k' \approx (1 - h(\rho) + \varepsilon)n - \log n$ already ρ-covers almost all vectors in $\mathbb{F}_2^n$ with high probability. Then, it adds $\approx \log n$ extra random basis vectors to $\mathcal{C}'$ one-by-one and tracks the evolution of a (different) potential over these $\log n$ steps. The potential Blinovsky uses corresponds to the probability that a random $y \in \mathbb{F}_2^n$ is *remote*, meaning that it is covered by much fewer than the expected number of codewords in $\mathcal{C}'$, which is $\frac{|\mathcal{C}'| \cdot |B_{\rho n}(0)|}{2^n}$.

  In the end, we conclude that, with high probability, there are no remote points with respect to the final code $\mathcal{C}$. This implies not only that $\mathcal{C}$ is ρ-covering with high probability, but also the stronger lower bound

$$|\mathcal{C} \cap B_{\rho n}(y)| \geq \frac{1 - o(1)}{n} \cdot \frac{|\mathcal{C}| \cdot |B_{\rho n}(0)|}{2^n}$$

  for all $y \in \mathbb{F}_2^n$. This is a $1/n$ factor off the optimal bound. Recently, the stronger bounds

$$(1 - o(1)) \frac{|\mathcal{C}| \cdot |B_{\rho n}(0)|}{2^n} \leq |\mathcal{C} \cap B_{\rho n}(y)| \leq (1 + o(1)) \frac{|\mathcal{C}| \cdot |B_{\rho n}(0)|}{2^n},$$

  which we may see as a sharp list-decoding guarantee, were shown to hold for all y with high probability for a random linear code [DLM+26] (in fact, a more general result holds). This means that random linear codes match the behavior of uniformly random unstructured codes in list-decoding above capacity.

Follow-up work by Blinovsky [Bli90a] gives better bounds on the probability that a random linear code of rate $R > 1 - h(\rho)$ is ρ-covering.

- 1990: Blinovsky [Bli90b] extends the study of linear covering codes to covering by translates of a shape other than a Hamming ball. Roughly speaking, the ideal goal would be to show that for any set $S \subseteq \mathbb{F}_2^n$, a random linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$ of rate $R \approx 1 - \frac{1}{n} \log |S|$ will satisfy

$$\mathcal{C} + S = \mathbb{F}_2^n$$

  with high probability. Instantiating $S = B_{\rho n}(0)$ recovers the "ρ-covering problem" discussed above (see Equation (1)). Blinovsky establishes this result for "puffed-up" sets of the form $S' = S + B_\Delta(0)$ with S an arbitrary set and $\Delta = o(n)$.

- 1997: Alon, Dietzfelbinger, Miltersen, Petrank, and Tardós [ADM+97] use the potential method for a different covering-type problem, with the same potential. A connection between the potential method used in this work and that of GHSZ, in line with our comparison between Goblick's result and GHSZ above, was also remarked in Kumar's recent PhD thesis [Kum26, Section 3.1.2].

- 2000: GHSZ [GHSZ02] use the potential method with a new potential to show that linear codes achieve list-decoding capacity.

AI usage declaration. GPT-6 was used to critique the exposition and check for technical correctness in a first human-made full draft, and to aid the literature search. In particular, GPT-6 identified Levitin's work [Lev88] as relevant, suggested the unified interpretation of the potential method in the works of Goblick and GHSZ (which was incorporated above), and helped translate [Bli90a, Bli90b] from their original Russian and parse the techniques and results in these works.

References

[ADM+97] Noga Alon, Martin Dietzfelbinger, Peter Bro Miltersen, Erez Petrank, and Gábor Tardos. Is linear hashing good? In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC 1997)*, pages 465–474. ACM, 1997.

[Ber71] Toby Berger. *Rate Distortion Theory*. Prentice-Hall, Inc., 1971.

[Bli87] Vladimir Markovich Blinovskii. Lower asymptotic bound on the number of linear code words in a sphere of given radius in $\mathbb{F}_q^n$. *Problemy Peredachi Informatsii*, 23(2):50–53, 1987. Translated in: Problems of Information Transmission, vol. 23, no. 2, pp. 130–132.

[Bli90a] Vladimir Markovich Blinovskii. Asymptotically exact uniform bounds for spectra of cosets of linear codes. *Problemy Peredachi Informatsii*, 26(1):99–103, 1990. Translated in: Problems of Information Transmission, vol. 26, no. 1, pp. 83–86, 1990.

[Bli90b] Vladimir Markovich Blinovskii. Covering the Hamming space with sets translated by vectors of a linear code. *Problemy Peredachi Informatsii*, 26(3):21–26, 1990. Translated in: Problems of Information Transmission, vol. 26, no. 3, pp. 196–201, 1990.

[CF85] Gérard Cohen and Peter Frankl. Good coverings of Hamming spaces with spheres. *Discrete Mathematics*, 56(2):125–131, 1985.

[CG90] John T. Coffey and Rodney M. Goodman. The complexity of information set decoding. *IEEE Transactions on Information Theory*, 36(5):1031–1037, 1990.

[CHLL97] Gérard Cohen, Iiro Honkala, Simon Litsyn, and Antoine Lobstein. *Covering Codes*, volume 54 of *North-Holland Mathematical Library*. Elsevier, 1997.

[Coh83] Gérard Cohen. A nonconstructive upper bound on covering radius. *IEEE Transactions on Information Theory*, 29(3):352–353, 1983.

[DLM+26] Dean Doron, Tal Leonov, Jonathan Mosheiff, Henrique Navas, Nicolas Resch, and João Ribeiro. Discrepancy for random linear codes, 2026. https://arxiv.org/abs/2606.24471.

[DP86] Philippe Delsarte and Philippe Piret. Do most binary linear codes achieve the Goblick bound on the covering radius? (corresp.). *IEEE Transactions on Information Theory*, 32(6):826–828, 1986.

[GHSZ02] Venkatesan Guruswami, Johan Håstad, Madhu Sudan, and David Zuckerman. Combinatorial bounds for list decoding. *IEEE Transactions on Information Theory*, 48(5):1021–1034, 2002. Preliminary version presented at Allerton 2000.

[Gob62] Thomas J. Goblick. *Coding for a discrete information source with a distortion measure*. PhD thesis, MIT, October 1962. Available at https://dspace.mit.edu/entities/publication/fbba9539-8ae5-4d8a-bf89-6fb07116288a.

[Kum26] Vinayak M. Kumar. *Linearity, locality, and lower bounds: A unified perspective through randomness*. PhD thesis, UT Austin, 2026. https://v-m-kumar.github.io/dissertation%20(2).pdf.

[Lev88] Lev B. Levitin. Covering radius of almost all linear codes satisfies the Goblick bound, 1988. Presented at a recent results session, 1988 IEEE International Symposium on Information Theory.