

Quantum multi-anomaly detection

Santiago Llorens¹, Gael Sentís^{1,2}, and Ramon Muñoz-Tapia¹

¹Física Teòrica: Informació i Fenòmens Quàntics, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

²Ideated, Carrer de la Tecnologia, 35, 08840 Viladecans, Barcelona, Spain

A source assumed to prepare a specified reference state sometimes prepares an anomalous one. We address the task of identifying these anomalous states in a series of n preparations with k anomalies. We analyze the minimum-error protocol and the zero-error (unambiguous) protocol and obtain closed expressions for the success probability when both reference and anomalous states are known to the observer and anomalies can appear equally likely in any position of the preparation series. We find the solution using results from association schemes theory, thus establishing a connection between graph theory and quantum hypothesis testing. In particular, we use the Johnson association scheme which arises naturally from the Gram matrix of this problem. We also study the regime of large n and obtain the expression of the success probability that is non-vanishing. Finally, we address the case in which the observer is blind to the reference and the anomalous states. This scenario requires a universal protocol for which we prove that in the asymptotic limit, the success probability corresponds to the average of the known state scenario.

1 Introduction.

Over the past few decades, quantum information theory and its direct application to quantum technologies have advanced rapidly. Numerous quantum tasks in quantum computing [1], quantum simulation [2], quantum sensing [3], quantum communications [4], or quantum machine

learning [5, 6] rely on the ability of sources to prepare series of identical quantum states. However, the practical implementation of such identical state preparation faces significant challenges due to noise and imperfections in quantum state generation. In this work, we consider a source that is intended to produce a series of identical states, but it occasionally produces different states. These anomalies can be seen as outliers in a quantum data series, and our task is to identify them. This identification can be thought of as an instance of quality control in production, allowing for the application of corrective measures, thereby optimizing resource usage and improving the scalability of quantum systems for technical applications. This task is also important for the quantum internet [7], as it can be regarded as a protocol to detect intruders in quantum communications: detecting anomalies in the message may identify which information-carrying system has been compromised.

We study the basic setting of this problem when only a fixed number of mispreparations, or anomalies, occur. We assume that the anomalous states are known and arbitrary, and it is also promised that a given number of them have been produced, but their positions in the string of prepared states are not known. We aim to certify which of these positions correspond to anomalous states with the minimum error. Even though our results are derived under these assumptions, they immediately provide a benchmark for multi-anomaly detection strategies in more realistic scenarios, where one should not expect to know the anomalous states, nor these to be pure. Indeed, such generalizations correspond to scenarios where we have less information about the source, hence the success probability of correctly identifying the position of the anomalies will be upper bounded by that of the case of known and pure states. Partial results for the simplest case of one anomaly were presented in [8]. A somewhat

Santiago Llorens: santiago.llorens@uab.cat

Gael Sentís: gael.sentis@uab.cat

Ramon Muñoz-Tapia: ramon.munoz@uab.cat

different problem of detecting anomalies from a machine learning perspective was also discussed in [9].

The problem we are dealing with turns out to be an instance of multi-hypothesis discrimination, for which generically no analytical solutions can be found. Complete solutions are basically only known for the two-state case [10], a result that marks the beginning of the quantum state discrimination field [11, 12, 13], and sources of symmetric states [14, 15]. Applications to classification problems have recently been addressed [16], as well as to the detection of Hamiltonian changes [17]. Asymptotic expressions for some nearly symmetric sources such as the change point [18] and a few exact zero-error cases are also known [19, 20]. The multi-anomaly detection task exhibits symmetries that enable us to completely solve the problem.

This work first aims to derive the optimal solution for the minimum error approach and to study its asymptotic behavior for a large number of preparations. Additionally, we address the zero error protocol, or unambiguous identification, in which no error is allowed when identifying the anomalies. We also find the optimal universal protocol [21] that arises when considering that both reference and anomalous states are unknown to the observer. This protocol seeks to provide a robust solution that is applicable across different quantum systems and experimental conditions.

This paper is structured as follows. We first introduce the problem and its Gram formulation and present some mathematical results from graph theory and association schemes that prove to be very suitable for our problem. This toolbox allows us to demonstrate the optimality of the square root measurement (SRM) in a straightforward way. We then present the optimal solution for the minimum error protocol and show that for a fixed number of anomalies k , its success probability is non-vanishing in the asymptotic regime of large number of systems n . We next address the unambiguous protocol and compare the results with the minimum error approach. Finally, we present an optimal protocol that is blind to the information of the reference and anomalous states. We end the paper with three technical appendices, containing a brief account of Johnson graphs and Bose-Mesner algebra, Schur-Weyl duality, and explicit success probability formulae.

2 Setting of the problem.

Let us start by describing the key elements of the multi-anomaly detection problem. We denote the reference state of the string as $|\phi_0\rangle$ and the anomalous state as $|\phi_1\rangle$. For a given size of the string n and a fixed number k of anomalous states, we have $N = \binom{n}{k}$ possible hypotheses to consider. Without loss of generality, we can consider the overlap between the reference and an anomalous state to be real and denote it as $c = |\langle\phi_0|\phi_1\rangle|$. We also assume that each hypothesis has the same prior probability to occur. We denote the global state of hypothesis r by $|\Psi_r\rangle$, with r a subset of $\{1, \dots, n\}$ with cardinality $|r| = k$ that labels the position of the anomalies (see Fig. 1 for simple examples).

We next construct the Gram matrix of the set of hypotheses, which contains all the relevant information for a discrimination task. It is defined as the matrix of overlaps of the hypotheses, which in our case reads

$$G_{rs} = \langle\Psi_r|\Psi_s\rangle = (c^2)^{\delta(r,s)}, \quad (1)$$

where $\delta(r, s) := k - |r \cap s|$ is defined as the distance between two subsets r and s , and where $|r \cap s|$ is the number of elements belonging to the intersection of the subsets r and s and corresponds to half the Hamming distance. In Fig. (1) we depict the cases of one and two anomalies for $n = 4$ systems. In the first case, the distance between different hypotheses is always $\delta = 1$, while in the second case, we have $\delta = 1$ and $\delta = 2$ distances.

Decomposing the Gram matrix G in powers of c^2 we obtain a collection of symmetric matrices, that we will denote by A_i , whose entries are 0's and 1's. Thus, we have

$$G = \sum_{i=0}^k (c^2)^i A_i. \quad (2)$$

In particular, A_i are the adjacency matrices of the generalized Johnson graphs, and form the so-called Bose-Mesner algebra \mathcal{A} of the Johnson association scheme [22]. The elements of this algebra commute and all have zeros in their diagonal entries, except for the identity matrix $\mathbb{1} = A_0$ (see Appendix A for more details). As we see below, standard results of this association scheme allow us to provide proof of optimality as well as an elegant method to obtain analytical expressions for the success probability.

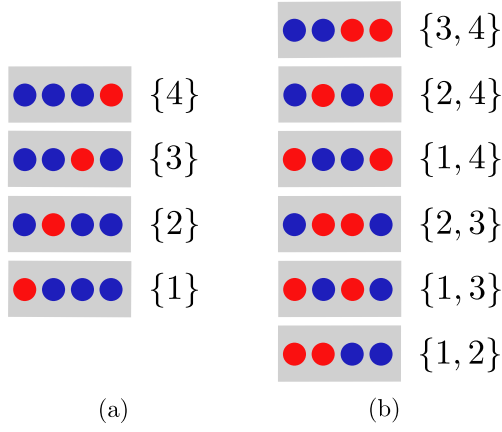


Figure 1: Set of hypotheses for the cases of $n = 4$. Subfigure (a) corresponds to $k = 1$, notice that the distance between any two hypotheses is $\delta = 1$. Subfigure (b) corresponds to $k = 2$ and distances can take values of $\delta = 1$, for example hypotheses $\{3, 4\}$ and $\{2, 4\}$ or $\delta = 2$, for instance hypotheses $\{3, 4\}$ and $\{1, 2\}$.

3 Minimum error.

The minimum error protocol consists of a measurement procedure given by a positive operator valued measure (POVM), i.e. a set of positive semidefinite operators $\{\Pi_r \geq 0\}$ that satisfy the completeness relation $\sum_r \Pi_r = \mathbb{1}$. To each element Π_r of the POVM one associates the hypothesis $|\Psi_r\rangle$. The conditional probability of obtaining outcome Π_s given a state $|\Psi_r\rangle$ is given by the Born rule $\text{Tr}(|\Psi_r\rangle\langle\Psi_r|\Pi_s)$ (we have made a slight abuse of notation and labeled the outcome by the same name of the POVM operator). Success occurs when, upon measuring $|\Psi_r\rangle$, one obtains outcome Π_r and then the average success probability reads

$$P_S = \sum_r \eta_r \text{Tr}(|\Psi_r\rangle\langle\Psi_r|\Pi_r), \quad (3)$$

where η_r is the prior probability of occurrence of hypothesis $|\Psi_r\rangle$. From now onward we refer to the average P_S simply as the success probability and assume that all hypotheses have equal priors, $\eta_r = \eta_s = 1/N = \binom{n}{k}^{-1} \forall r, s$.

The square root $\sqrt{G} = S$ of a Gram matrix defines a measurement, the so-called SRM [23], as the elements S_{rs} are the projections into a measurement basis $\{|m_k\rangle\}$ defined by S , i.e., $S_{rs} = \langle m_r | \Psi_s \rangle$. Therefore for equal priors, the average success probability of such measurement reads

$$P_S = \frac{1}{N} \sum_r |S_{rr}|^2. \quad (4)$$

Now, since the Gram matrix Eq. (2) belongs to a Bose-Mesner algebra \mathcal{A} , any function of G also belongs to the same algebra, and in particular its square root $S = \sqrt{G}$. Therefore, given that the diagonal entries of G are constant, so are the diagonal entries of S , that is, $S_{ll} = S_{hh} = \forall l, h$, because these only depend on the element of the algebra $A_0 = \mathbb{1}$. Optimality then follows directly, as it is known that for constant diagonal terms of S the SRM is optimal [18, 24].

We next write the success probability in a more convenient way. We note that since the diagonal terms of S are all equal, they can be expressed as $S_{ll} = (\text{Tr}S)/N$, with $\text{Tr}S = \sum_l \sqrt{\lambda_l}$, where λ_l are the eigenvalues of the Gram matrix. We obtain [18]

$$P_S = \left(\frac{\sum_{l=1}^N \sqrt{\lambda_l}}{N} \right)^2 \quad (5)$$

Hence, the job reduces to finding the eigenvalues λ_l . As every adjacency matrix A_i commutes with each other, they are all simultaneously diagonalizable in some basis, and therefore the eigenvalues of G can be obtained by summing the contributions of a given eigenvalue for each A_i

$$\lambda_j(G) = \sum_{i=0}^k (c^2)^i \lambda_j(A_i). \quad (6)$$

For the rest of this work, we will omit the dependence of the G eigenvalues and simply write λ_j for $\lambda_j(G)$. For the eigenvalues of the adjacency matrices, we will however write the explicit dependence.

From standard results of association schemes the eigenvalues of the first adjacency matrix $A := A_1$ determine the eigenvalues of any other adjacency matrix of the scheme. The eigenvalue $\lambda_j(A_i)$ is the value of a certain orthogonal polynomial evaluated at the eigenvalues of A [22]

$$\lambda_j(A_i) = (-1)^i \binom{k}{i} R_i(\lambda_j(A) + k; 0, n - 2k, k), \quad (7)$$

where R_i is the dual Hahn polynomial of degree i [25], which are the orthogonal polynomials arising from the Johnson scheme. Introducing Eq. (7) in Eq. (6) we note that the eigenvalues of G are the generating functions of the dual Hahn polynomials, which read

$$\lambda_j = (1 - c^2)^j {}_2F_1 \left(\begin{matrix} j - k, & -n + k + j \\ & 1 \end{matrix}; c^2 \right), \quad (8)$$

with multiplicities

$$m_j = \binom{n}{j} - \binom{n}{j-1}, \quad (9)$$

where $j = 0, \dots, k$, and ${}_2F_1$ is the hypergeometric function [26] (see Eq. (36) for further details). Note that the number of distinct eigenvalues is $k + 1$, i.e. just the number of anomalies plus one, and are monotonically decreasing with j . For instance, in the simplest case of one anomaly, there are only two different eigenvalues, $\lambda_0 = 1 + (n-1)c^2 > \lambda_1 = (1-c^2)$ with multiplicities 1 and $n-1$, respectively.

Taking into account the degeneracy, the success probability of the minimum error protocol finally reads

$$P_S = \left(\sum_{j=0}^k \frac{m_j}{N} \sqrt{\lambda_j} \right)^2, \quad (10)$$

where λ_j are given in Eq. (8).

For a fixed number of anomalies and an increasing number of total systems n , it could be expected that the success probability vanishes as the number of possible hypotheses goes to infinity (see Fig. 2). However, the asymptotic limit of P_S for $n \rightarrow \infty$ is finite as we next show. We note that the eigenvalues of the Gram matrix are polynomials in n of degree $k-j$ ($j = 0, 1, \dots, k$), i.e. $\lambda_j \sim O(n^{k-j})$. On the other hand, the ratio between the multiplicities of each eigenvalue m_j and the dimension of the space scales as $m_j/N \sim O(n^{j-k})$ (see Appendix C for the explicit expression of simple cases). Then, it follows that the terms in Eq. (10) contributing to the first two leading terms are proportional to λ_k and $\sqrt{\lambda_k \lambda_{k-1}}$. Substituting the values of the multiplicities, Eq. (9), and the explicit expression of the eigenvalues λ_k, λ_{k-1} given in Eq. (7) [listed in Table (3) of Appendix A], we finally get

$$P_S = (1-c^2)^k + \frac{2kc(1-c^2)^{k-1/2}}{\sqrt{n}} + O\left(\frac{1}{n}\right). \quad (11)$$

Note that this expansion is valid in the regime $k/n \ll 1$.

In Fig. 2 we depict the success probability for two and three anomalies as a function of the string length. As can be seen, the convergence to the asymptotic value is rather slow as the first correction is of order $O(1/\sqrt{n})$. The values of the

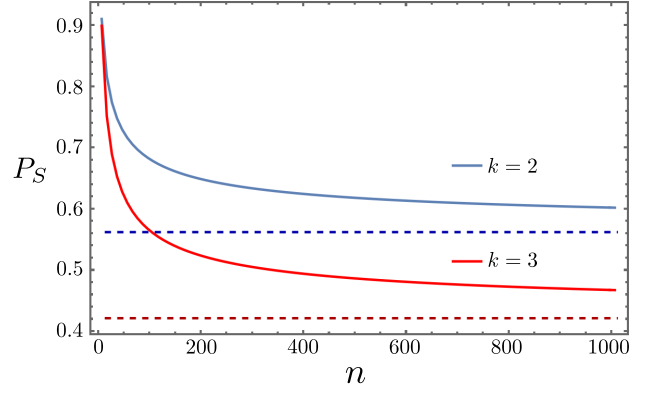


Figure 2: Comparison of the success probability of 2 anomalies, the upper curve (blue) and the case of 3 anomalies, the lower curve (red), both for a fixed overlap $c = \frac{1}{2}$. The dashed, upper and lower, lines correspond to the asymptotic limits of the success probabilities of 2 and 3 anomalies respectively.

success probability differ considerably from the asymptotic value, even for a sizeable value of n . This difference is also larger for a larger number of anomalies because of the k factor appearing in the second term of Eq. (11).

4 Zero error identification.

It is rather straightforward to tackle also the zero error identification scenario, also known as unambiguous discrimination, where one demands that the protocol provides outcomes that identify the position of the anomalies without error. Naturally, this can only happen at the expense of having some inconclusive outcomes.

The aim is to find a protocol that maximizes the rate of conclusive outcomes, or equivalently minimizes the rate of the inconclusive ones. The POVM has $N + 1$ elements: $\{\Pi_r\}$ with $r \subset \{1, \dots, n\}$ with $|r| = k$, and a Π_0 element which corresponds to the inconclusive outcome. Each Π_r element identifies with zero error the state $|\Psi_r\rangle$. The structure of the POVM operators is tightly constrained by the zero error conditions $\text{Tr}(\Pi_r |\Psi_s\rangle \langle \Psi_s|) = 0$ for $r \neq s$. The optimization can be written as a semidefinite program (SDP) [27, 28],

$$P_S = \frac{1}{N} \max_{\Gamma} \text{Tr } \Gamma \quad (12)$$

s.t. $G - \Gamma_D \geq 0$ and $\Gamma \geq 0$

Here the matrix Γ_D is the diagonal part of Γ i.e., $[\Gamma_D]_{rs} = \delta_{rs} \Gamma_{rr}$. The entries Γ_{rr} correspond to

the conditional probabilities that, given hypothesis $|\Psi_r\rangle$, it is detected without error. This form of the SDP is easily derived by noticing that a POVM satisfying the zero error conditions must have elements of the type $E_k = \gamma_k |\tilde{\varphi}_k\rangle\langle\tilde{\varphi}_k|$, where the (unnormalized) states satisfy $\langle\tilde{\varphi}_k|\Psi_l\rangle = \delta_{kl}$, i.e., these states are the rows of the (pseudo) inverse of the matrix $X = \sum_k |\Psi_k\rangle\langle k|$, where $|k\rangle$ is an orthonormal basis of dimension N . The condition $G - \Gamma_D$ follows from multiplying the positivity condition of the inconclusive POVM element, $\mathbb{1} - \sum_k \gamma_k |\tilde{\varphi}_k\rangle\langle\tilde{\varphi}_k| \geq 0$, left and right by the operators X^\dagger and X , respectively (see [20] for more details).

To obtain the solution of Eq. (12) we consider the ansatz

$$\Gamma = (1 - c^2)^k \mathbb{1}, \quad (13)$$

which yields a success probability

$$P_S = (1 - c^2)^k. \quad (14)$$

We prove that Eq. (14) is the optimal value by checking that (i) Γ belongs to the feasibility set (i.e. $G - \Gamma_D \geq 0$ and $\Gamma \geq 0$) and that (ii) it induces a feasible instance of the dual SDP in Eq. (12). Property (i) follows directly from the fact that $(1 - c^2)^k = \lambda_{\min}$ (see Table 2). For the second property let us write the dual program [20]

$$P_S = \frac{1}{N} \min_Y \text{Tr } GY, \quad (15)$$

s.t. $Y_D - \mathbb{1} \geq 0$ and $Y \geq 0$,

and let us choose

$$Y = \frac{N}{m_k} E_k, \quad (16)$$

where E_k is the projector onto the minimal eigenvalue of G [see Eq. (30) in Appendix A]. From this choice, it follows that $P = \text{Tr } GY$ coincides with the value of the primal SDP. So, if Y is a feasible solution to Eq. (15), it is an optimal solution as the values of both primal and dual objective functions coincide. We just have to prove that Y_D , the matrix of the diagonal terms of Y , satisfies $Y_D \geq \mathbb{1}$ (note that Y is non-negative by construction as E_k is an orthogonal projection). Let us decompose E_k as a linear combination of adjacency matrices, $E_k = (1/N) \sum_{j=0}^k q_k(j) A_j$, where $q_k(j)$ are given by the Hahn polynomials [see Eqs (32) in Appendix A] and note again that

the only matrix with diagonal elements different from 0 is $A_0 = \mathbb{1}$, i.e.,

$$Y_D = \frac{q_k(0)}{m_k} \mathbb{1}. \quad (17)$$

But $q_k(0) = m_k$ [see Eq. (33b) and Table 1] and hence $Y_D = \mathbb{1}$, which completes the proof.

This result could have been readily anticipated from (i) as the symmetry of the problem already requires that the optimal Γ_D satisfies $\Gamma_D \propto \mathbb{1}$ and then Eq. (12) is just the program defining the lowest eigenvalue of G . We also see that Eq. (14) coincides with the leading term of the minimum error success probability, Eq. (11), hence we conclude that the unambiguous measurement is also optimal for the minimum error protocol for large n . This measurement can be realized by a local protocol that checks if each particle has projection or not in the complementary subspace of the reference state. This result can be extended to the case where the anomalies are mixed states since the optimal protocol is fixed by the reference state.

5 Universal protocol

. We finally tackle the anomaly identification task when the reference and anomalous states are unknown to the observer. That is, we aim at finding a universal protocol that does not use any information about the states (only that they are different, of course) and hence works for an arbitrary pair of reference and anomalous states. In this section, we closely follow the setup and notation of Ref. [16].

In this scenario, all hypotheses are generated by applying a permutation to the fiducial state, in which all k anomalies are located at the end of the string

$$\rho_\sigma = \int U_\sigma [\phi_0]^{\otimes(n-k)} \otimes [\phi_1]^{\otimes k} U_\sigma^\dagger d\phi_0 d\phi_1, \quad (18)$$

where $[\cdot] := |\cdot\rangle\langle\cdot|$, ϕ_0 (ϕ_1) corresponds to the reference (anomalous) state, U_σ stands for the unitary transformation that acts on the tensor-product space as a permutation $\sigma \in \mathcal{S}_n \subset S_n$, where $d\phi_i$ is the measure of the uniform distribution of $|\phi_i\rangle$ and S_n is the symmetric group of n elements. Here, \mathcal{S}_n is the subgroup of relevant permutations for k anomalies, that is, permutations yielding the same state ρ_σ are accounted only once. It is

convenient to index each hypothesis by the permutation σ that takes a fiducial state where all anomalies occur in the last positions of the hypothesis in question. Using Schur lemma [29, 16] (see Appendix B), the hypotheses read

$$\rho_\sigma = c_k U_\sigma \mathbb{1}_{n-k}^{\text{sym}} \otimes \mathbb{1}_k^{\text{sym}} U_\sigma^\dagger, \quad (19)$$

where $\mathbb{1}_k^{\text{sym}}$ refers to the projector onto the fully symmetric subspace of k parties. In the Schur basis, it reads

$$\rho_\sigma = c_k \bigoplus_\lambda \mathbb{1}_\lambda \otimes \Omega_\sigma^\lambda, \quad (20)$$

where the normalization constant c_k is

$$c_k = \frac{1}{d_k^{\text{sym}} d_{n-k}^{\text{sym}}}, \quad (21)$$

and

$$d_k^{\text{sym}} = \binom{k+d-1}{d-1} \quad (22)$$

is the dimension of the symmetric subspace of k parties. The parameter λ labels the irreducible representation (irrep for short) of the joint action of the groups $\text{SU}(d)$ and the symmetric group S_n over the whole state space $(d, \mathbb{C})^{\otimes n}$, and it is usually identified with bipartitions $(n-\lambda, \lambda)$ of n elements. The identity $\mathbb{1}_\lambda$ is defined over the subspace associated with the irrep λ of $\text{SU}(d)$, and arises in Eq. (20) because we are averaging over all possible states. The second factor Ω_σ^λ stands for an operator that acts in the irrep of S_n labeled by λ and it is a rank-1 projector [16].

We can now derive the explicit expression of the success probability for the minimum-error protocol. The optimal measurement, fortunately, turns out to be again the SRM. For mixed states, its elements are $\Pi_\sigma = \rho^{-1/2} \rho_\sigma \rho^{-1/2}$, where

$$\rho = \sum_{\sigma \in S_n} \rho_\sigma = \bigoplus_\lambda C_\lambda \mathbb{1}_\lambda \otimes \mathbb{1}^\lambda. \quad (23)$$

where we have used Schur lemma, since ρ commutes with every element $\sigma \in S_n$, and C_λ is a proportionality constant for each partition λ . As in previous sections, we write the dual SDP for the success probability: $\min_Y \text{Tr } Y$, such that $Y - \rho_\sigma \geq 0$ for all ρ_σ . If we consider $Y = \sum_\sigma \Pi_\sigma \rho_\sigma$, primal and dual programs give the same value. To prove optimality, one only has to check that Y satisfies the feasibility conditions. Naturally, $Y - \rho_\sigma \geq 0$, are just the well-known Holevo conditions [30]. It is easy to compute that $Y = c_k \bigoplus_\lambda \mathbb{1}_\lambda \otimes \mathbb{1}^\lambda$, and to check that

$Y - \rho_\sigma \geq 0$, $\forall \sigma \in S_n$, holds, since $\mathbb{1}^\lambda \geq \Omega_\sigma^\lambda$. Hence, the optimal success probability of the universal protocol is just $(1/N) \text{Tr } Y$, which reads

$$P_S = \sum_{\lambda=0}^k \left(\frac{(n-2\lambda+1)^2}{(n-\lambda+1)^2} \frac{\binom{n-\lambda+d-1}{d-1}}{\binom{n-k+d-1}{d-1}} \right) \times \frac{\binom{n}{\lambda} \binom{\lambda+d-2}{d-2}}{\binom{n}{k} \binom{k+d-1}{d-1}}, \quad (24)$$

where we have used the explicit formulae for the dimensions of the irreps of $\text{SU}(d)$ and S_n for a bipartition $(n-\lambda, \lambda)$ [see Eqs. (42a, 42b)].

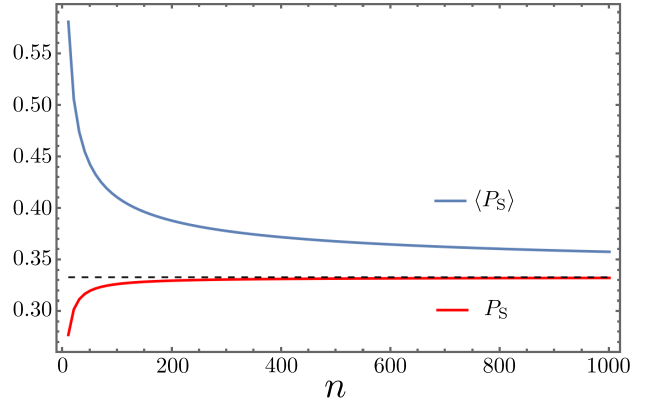


Figure 3: Comparison between the average (over c) success probability for the case of known reference and anomalous states of dimension $d = 2$, the upper curve (blue online), and the success probability for the universal protocol, the lower curve (red online). The dashed black line corresponds to the asymptotic limit of both success probabilities $P_S = 1/(k+1)$

The asymptotic behavior of the success probability in the regime of large n is dominated by the partition $\lambda = k$, as the most important factor in Eq. (24) is the ratio $\binom{n}{\lambda} / \binom{n}{k}$ and the contributions of the rest of partitions $\lambda = l < k$ vanish as $O(1/n^{k-l})$. The asymptotic expansion of the success probability then reads

$$P_S = \frac{d-1}{d-1+k} + O\left(\frac{1}{n}\right) \quad (25)$$

Interestingly, this value coincides with the asymptotic limit of the average success probability for known states and anomalies (see Fig. 3),

$$\begin{aligned} \langle P_S \rangle &\sim \int_0^1 (1-c^2)^k (1-c^2)^{(d-2)} (d-1) dc^2 \\ &= \frac{d-1}{d-1+k} \end{aligned} \quad (26)$$

where $(d-1)(1-c^2)^{(d-2)} dc^2 =: d\mu(c^2)$ is the uniform measure of the square of the overlap [16] and

we have dropped the vanishing terms in Eq. (11). This agreement has also been observed in the change point problem for unknown states [31]. The reasons behind this coincidence here stem from the fact that, for known states and large n , the optimal measurement approaches an unambiguous protocol that detects anomalies by projecting on the orthogonal space of the reference state. Therefore, the knowledge of the anomalous state is not required, for it would be impossible to determine this state with a fixed (and small) number of instances distributed at random places. In this asymptotic limit, only the reference state needs to be known to devise the protocol, and this can be done by using a vanishing number of systems (e.g. $\propto \sqrt{n}$). The success probability Eq. (24) is depicted in Fig. 3. We observe that, as n grows, somehow counterintuitively, P_S increases. The logic that explains this behavior is that increasing n allows the protocol to gain more knowledge about the reference state, which gives an advantage that exceeds the detrimental effect of having a larger set of possible hypotheses. It is also evident that, although the universal protocol reaches the value Eq. (26) rather fast, the average of the success probability for known reference states and anomalies is notably above this value even for significantly large string lengths. Note also that, as expected, $\langle P_S \rangle$ in Eq. (26) tends to one as $d \rightarrow \infty$, which reflects the fact that random states tend to be more orthogonal as d grows.

6 Conclusions.

We have done a comprehensive analysis of the problem of identifying anomalous states in a series of allegedly identical quantum state preparations. The accurate identification of faulty preparations will be crucial in the development of real quantum applications. Our results provide a theoretical benchmark of performance under some simplifying assumptions. A key ingredient to finding closed analytical expressions for the success probability is the theory of association schemes. This mathematical area finds a rele-

vant physical realization in our problem. We have proved that, in the asymptotic limit of large n , both optimal minimum error and zero error identification tasks use the same measurement strategy. Furthermore, the success probability is finite and decreases as a power of the number of anomalies. We have also presented a universal protocol that is independent of the states appearing in the problem. Results along these lines have been obtained in the context of quantum learning [16, 21] using representation theory and the Schur-Weyl decomposition. We have also analyzed the performance of this protocol in the asymptotic limit. We obtain that the success probability coincides with the average of the success probability for known states. This result is quite unexpected in this context since the number of anomalous states is small and these states are located in random places, but it can be explained from the structure of the optimal measurement in this limit.

The tools presented here can be applied to other problems from a novel perspective, e.g., unsupervised classification of states using other figures of merit beyond the error probability. In general, we expect association schemes such as Hamming and Johnson schemes along with the ones arising from cyclic graphs to be useful for problems in quantum information with appropriate underlying symmetries.

7 Acknowledgements.

This work has been financially supported by Ministerio de Ciencia e Innovación of the Spanish Government with funding from European Union Next GenerationEU (PRTR-C17.I1) and by the Generalitat of Catalunya, by the Ministry of Economic Affairs and Digital Transformation of the Spanish Government through the QUANTUM ENIA project: Quantum Spain, by the European Union through the Recovery, Transformation and Resilience Plan - NextGenerationEU within the framework of the “Digital Spain 2026 Agenda”, and from the Spanish Agencia Estatal de Investigación, Grants Nos. PID2019-107609GB-I00 and PID2022-141283NB-I00.

A Johnson scheme

In this appendix, we introduce the necessary tools of association scheme theory [22] for the procedure when solving the multi-anomaly detection problem. First, we will define some necessary concepts,

and later we will relate them to our problem. This will allow us to prove optimality conditions for unambiguous discrimination or to compute the values of the success probability of minimum error protocol.

Definition 1. A Johnson graph $J(n, k)$ is a graph whose vertices are the k -element subsets of a n -element set. Any pair of vertices r, s are connected by an edge if the intersection of the subsets $|r \cap s| = 1$ (see Figure 4).

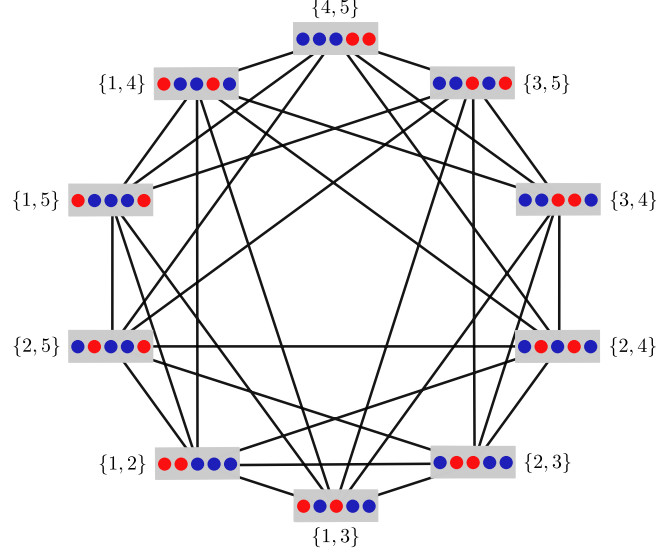


Figure 4: Johnson graph $J(5, 2)$ for a set of $n = 5$ objects with subsets of cardinality $k = 2$.

Definition 2. An adjacency matrix of a graph is a $(0,1)$ -matrix whose entries are defined as

$$A_{rs} = \begin{cases} 1 & \text{if vertices } r, s \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

The $k + 1$ eigenvalues and multiplicities of the Johnson graph $J(n, k)$ adjacency matrix are given by

$$\lambda_j(A) = j(n - k - j + 1) - k, \quad (28a)$$

$$m_j(A) = \binom{n}{k-j} - \binom{n}{k-j-1}, \quad (28b)$$

for $j = 0, \dots, k$.

Definition 3. A generalized Johnson graph $J(n, k, i)$ is a graph whose vertices are the k -element subsets of a n -element set. Any pair of vertices r, s are connected by an edge if the intersection of the subsets $|r \cap s| = i$.

Definition 4. A symmetric association scheme is a set of Boolean matrices $\{A_i\}_{i=0}^k$ which satisfy

$$\text{i) } \sum_{i=0}^k A_i = J \text{ (the all-ones matrix),} \quad (29a)$$

$$\text{ii) } A_0 = \mathbb{1}, \quad (29b)$$

$$\text{iii) } A_i A_j = A_j A_i = \sum_k p_{ij}^k A_k, \quad (29c)$$

for some constants p_{ij}^k called the intersection numbers of the scheme.

The adjacency matrices of generalized Johnson graphs form the so-called Johnson association scheme. Also, these matrices span a space denoted by \mathcal{A} of dimension $k + 1$, which from Eq. (29c) is closed under multiplication (commutative), that is, it is a commutative algebra that we call Bose-Mesner algebra.

Since all matrices belonging to \mathcal{A} commute, they are simultaneously diagonalizable in a certain basis, and therefore, \mathcal{A} has a unique basis of projectors $\{E_i\}_{i=0}^k$ satisfying

$$\text{i)} \quad E_i E_j = \delta_{ij} E_i, \quad (30a)$$

$$\text{ii)} \quad E_0 = \frac{1}{N} J, \quad (30b)$$

$$\text{iii)} \quad \sum_{i=0}^k E_i = \mathbb{1}, \quad (30c)$$

where $N = \binom{n}{k}$, in particular, the adjacency matrices of generalized Johnson graphs can be written on this basis

$$A_j = \sum_i p_j(i) E_i, \quad (31)$$

and vice versa

$$E_j = \frac{1}{N} \sum_i q_j(i) A_i, \quad (32)$$

where, $p_j(i)$ ($q_j(i)$) are the entries of the so-called eigenmatrix P (Q) of the scheme, which read

$$p_i(j) = k_i Q_j(i; -n + k - 1, -k + 1, k), \quad (33a)$$

$$q_j(i) = m_j Q_j(i; -n + k - 1, -k + 1, k), \quad (33b)$$

where $k_i = \binom{k}{i} \binom{n-k}{i}$ is the valency (number of neighbors) in the Johnson graph $J(n, k, i)$, m_j is the multiplicity of the eigenvalue j in Eq. (28a) and Q_j stands for the Hahn polynomials [25] of degree j . In Table 1, we show the first three Hahn polynomials.

j	$Q_j(x; -n + k - 1, -k - 1, k)$
0	1
1	$1 - \frac{nx}{k(n-k)}$
2	$1 - \frac{2(n-1)x}{k(n-k)} + \frac{(n-1)(n-2)x(x-1)}{k(k-1)(n-k)(n-k-1)}$

Table 1: First Hahn polynomials of degree j for arbitrary values of n and k .

For a symmetric association scheme, we can write any matrix A_i as a polynomial $v_i(A_1)$, with A_i as the adjacency matrix of the Johnson graph $J(n, k, i)$ and v_i some polynomial of degree i . For the Johnson association scheme, we have

$$A_i = (-1)^i \binom{k}{i} R_i(A + k; 0, n - 2k, k), \quad (34)$$

with R_i the dual Hahn polynomial of degree i . These polynomials read

$$\begin{aligned} R_i(\lambda(x); \gamma, \delta, N) \\ = {}_3F_2 \left(\begin{matrix} -i, & -x, & x + \gamma + \delta + 1 \\ & \gamma + 1, & -M \end{matrix} ; 1 \right), \end{aligned} \quad (35)$$

where $\lambda(x) = x(x + \gamma + \delta + 1)$ and the hypergeometric function is defined as

$${}_pF_q\left(\begin{matrix} a_1, & \dots, & a_p \\ b_1, & \dots, & b_q \end{matrix}; z\right) = \sum_{m=0}^{\infty} \frac{(a_1)_m \dots (a_p)_m}{(b_1)_m \dots (b_q)_m} \frac{z^m}{m!}, \quad (36)$$

where $(a)_m$ is the rising Pochhammer symbol [26], $(a)_m := a(a+1)\dots(a+m)$ and $(a)_0 := 1$. In the Johnson scheme, the values of the variables are given by $x = j = 0, \dots, k$, $\gamma = 0$, $\delta = n - 2k$ and $M = k$ (see Table 2 for the values of the first polynomials). Notice from Eqs. (33a, 34) the duality between the Hahn and the dual Hahn polynomials, as we have written the entries of the eigenmatrix P as the Hahn polynomials, but we could have written them as its dual counterpart, $R_i(\lambda(j); \gamma, \delta, M) = Q_j(i; \gamma, \delta, M)$.

i	$R_i(\lambda(x); 0, n - 2k, k)$
0	1
1	$1 - \frac{x(n - x + 1)}{k(n - k)}$
2	$1 - \frac{2x(n - x + 1)}{k(n - k)} + \frac{(x - 1)x(n - x + 1)(n - x)}{(k - 1)k(n - k)(n - k - 1)}$

Table 2: First dual Hahn polynomials of degree i for arbitrary values of n and k .

One of the generating functions of the dual Hahn polynomials is given by [25]

$$\begin{aligned} & \sum_{i=0}^k (-c^2)^i \binom{k}{i} R_i(\lambda_j + k; 0, n - 2k, k) \\ &= (1 - c^2)^{k-j} {}_2F_1\left(\begin{matrix} -j, & -n + 2k - j \\ 1 \end{matrix}; c^2\right), \end{aligned} \quad (37)$$

and, as pointed out in the main text, this expression (37), corresponds to the eigenvalues of the Gram matrix (6), which, for the sake of clarity, are presented in Table 3.

λ_k	$(1 - c^2)^k$
λ_{k-1}	$(1 - c^2)^{k-1} [1 + c^2(n + 1 - 2k)]$
λ_{k-2}	$(1 - c^2)^{k-2} \left[1 + \frac{1}{2}c^2(4 + c^2(n + 1 - 2k))(n + 2 - 2k)\right]$
\vdots	\vdots
λ_0	$\sum_{i=0}^k c^{2i} \binom{k}{i} \binom{n - k}{i}$

Table 3: First two smallest and maximum eigenvalues of a Gram matrix with arbitrary parameters n, k, c .

B Schur-Weyl duality

In this appendix, we recall a brief introduction to representation theory used for the universal protocol of multi-anomaly detection with a connection of representation theory with the Johnson scheme.

Schur-Weyl duality establishes a connection between the irreducible representations of the group of linear transformations $\text{GL}(d)$, in particular for our problem, $\text{SU}(d)$, and the symmetric group S_n . The action of any transformation $R^{\otimes n}$ with $R \in \text{SU}(d)$, commutes with the action of any U^σ , the permutation of the tensor-product space $(d, \mathbb{C})^{\otimes n}$, with $\sigma \in S_n$. Both of these transformations induce

a reducible representation over the $SU(d)$ and S_n groups, and it follows that this representation reduces into irreps. as

$$U_\sigma R^{\otimes n} = R^{\otimes n} U_\sigma = \bigoplus_\lambda R_\lambda \otimes U_\sigma^\lambda, \quad (38)$$

where λ is a partition of the tensor-product space $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s, \dots)$, with $\sum_r \lambda_r = n$, that labels both, the irreducible representations of $R^{\otimes n}$ and U_σ . Note that our notation uses a subscript λ when referring to the representation of the $SU(d)$ group and a superscript when referring to S_n .

This block-diagonal structure allows us to decompose the total Hilbert space $\mathcal{H}^{\otimes n} = (d, \mathbb{C})^{\otimes n}$ as a direct sum of subspaces $\mathcal{H}^{\otimes n} = \bigoplus_\lambda H_\lambda \otimes H^\lambda$ that are invariant under the action of the groups $SU(d)$ and S_n . The basis in which the Hilbert space has this form is called the Schur basis, and it happens to be very suitable for universal protocols in which no information about the representation of the states in $SU(d)$ is available.

In our problem, we make use of this duality, since the only available information contained in our hypotheses is given by the representation of the S_n group. We start from Eq. (18) where we average over all possible uniform distributions of ϕ_0 and ϕ_1 . Then, using Schur lemma [29, 16], the hypotheses read

$$\rho_\sigma = c_k U_\sigma \mathbb{1}_{n-k}^{\text{sym}} \otimes \mathbb{1}_k^{\text{sym}} U_\sigma^\dagger, \quad (39)$$

where $\mathbb{1}_k^{\text{sym}}$ refers to the symmetric projector onto k parties, The normalization constant c_k in Eq. (21) is the inverse of the dimensions of the symmetric subspaces of k and $n - k$ parties

$$c_k = \frac{1}{d_k^{\text{sym}} d_{n-k}^{\text{sym}}}. \quad (40)$$

The dimension of the symmetric subspace of n parties corresponds to the dimension of the irrep. of $SU(d)$ of a bipartition $\lambda = (n, 0)$

$$d_k^{\text{sym}} = s_{(k,0)}. \quad (41)$$

the expression for $s_{(\lambda_1, \lambda_2)}$ is given explicitly in Eq. (42a).

States in (19) written using Schur basis [29] have a final block-diagonal form of Eq. (20). Notice that $\mathbb{1}_\lambda$ appears as the irrep. of $SU(d)$ in all bipartitions λ , since there is no information left of the possible directions that both reference and anomalous states can have.

Note that since, we are considering only two different states, the reference and the anomalous, the partitions of the tensor-product space will be 2-partitions, i.e. $\lambda = (\lambda_1, \lambda_2)$, with $\lambda_1 + \lambda_2 = n$. With this at hand, the dimensions of the irreps. $SU(d)$ and S_n used in Eq. (24) have a rather simple expression

$$s_\lambda = \frac{\lambda_1 - \lambda_2 + 1}{\lambda_1 + 1} \binom{\lambda_1 + d - 1}{d - 1} \binom{\lambda_2 + d - 2}{d - 2}, \quad (42a)$$

$$m_\lambda = \binom{n}{\lambda_2} - \binom{n}{\lambda_2 - 1}, \quad (42b)$$

notice that the label m in Eq. (42b) for the dimension of the S_n irrep. is not coincidental, as it corresponds to the multiplicity of the eigenvalues of the Jonshon graph in Eq. (28b).

C Success probability for 1, 2, and 3 anomalies

We present the explicit expressions of the success probability for minimum error P_S for the cases of $k = 1, 2, 3$ anomalies.

$$k = 1$$

$$P_S^{(1)} = \frac{1}{n^2} \left\{ (n-1) (1-c^2)^{1/2} + \sqrt{1+(n-1)c^2} \right\}^2. \quad (43)$$

$$k = 2$$

$$P_S^{(2)} = \left\{ \frac{n-3}{n-1} (1-c^2) + \frac{2}{n} (1-c^2)^{1/2} \sqrt{1+(n-3)c^2} + \frac{2}{n(n-1)} \sqrt{1+2\binom{n-2}{1}c^2 + \binom{n-3}{2}c^4} \right\}^2. \quad (44)$$

$$k = 3$$

$$\begin{aligned} P_S^{(3)} = & \left\{ \frac{n-5}{n-3} (1-c^2)^{3/2} + \frac{3(n-3)}{(n-1)(n-2)} (1-c^2) \sqrt{1+c^2(n-5)} \right. \\ & + \frac{6}{n(n-2)} (1-c^2)^{1/2} \sqrt{1+2c^2\binom{n-4}{1} + c^2\binom{n-5}{2}} \\ & \left. + \frac{6}{n(n-1)(n-2)} \sqrt{1+3c^2\binom{n-3}{1} + 3c^4\binom{n-4}{2} + c^6\binom{n-5}{3}} \right\}^2. \quad (45) \end{aligned}$$

References

- [1] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. *Nature* **574**, 505–510 (2019).
- [2] I.M. Georgescu, S. Ashhab, and Franco Nori. “Quantum simulation”. *Rev. Mod. Phys.* **86**, 153–185 (2014).
- [3] C.L. Degen, F. Reinhard, and P. Cappellaro. “Quantum sensing”. *Rev. Mod. Phys.* **89**, 035002 (2017).
- [4] Sumeet Khatri and Mark M. Wilde. “Principles of Quantum Communication Theory: A Modern Approach” (2024). [arXiv:2011.04672](#).
- [5] Vedran Dunjko and Hans J. Briegel. “Machine learning & artificial intelligence in the quantum domain: a review of recent progress”. *Rep. Prog. Phys.* **81**, 074001 (2018).
- [6] David Peral-García, Juan Cruz-Benito, and Francisco José García-Peñalvo. “Systematic literature review: Quantum machine learning and its applications”. *Comput. Sci. Rev.* **51**, 100619 (2024).
- [7] H Jeff Kimble. “The quantum internet”. *Nature* **453**, 1023–1030 (2008).
- [8] M. Skotiniotis, S. Llorens, R. Hotz, J. Calsamiglia, and R. Muñoz Tapia. “Identification of malfunctioning quantum devices” (2024). [arXiv:1808.02729](#).
- [9] Nana Liu and Patrick Rebentrost. “Quantum machine learning for quantum anomaly detection”. *Physical Review A* **97**, 042315 (2018).
- [10] Carl W. Helstrom. “Quantum detection and estimation theory”. *Journal of Statistical Physics* **1**, 231–252 (1969).
- [11] Anthony Chefles. “Quantum state discrimination”. *Contemporary Physics* **41**, 401–424 (2000).
- [12] Stephen M. Barnett and Sarah Croke. “Quantum state discrimination”. *Advances in Optics and Photonics* **1**, 238–278 (2009).
- [13] Joonwoo Bae and Leong-Chuan Kwek. “Quantum state discrimination and its applications”. *Journal of Physics A: Mathematical and Theoretical* **48**, 083001 (2015).

- [14] Stephen M. Barnett. “Minimum-error discrimination between multiply symmetric states”. *Physical Review A* **64**, 030303 (2001).
- [15] Y.C. Eldar, A. Megretski, and G.C. Verghese. “Optimal detection of symmetric mixed quantum states”. *IEEE Transactions on Information Theory* **50**, 1198–1207 (2004).
- [16] Gael Sentís, Alex Monràs, Ramon Muñoz-Tapia, John Calsamiglia, and Emilio Bagan. “Unsupervised Classification of Quantum Data”. *Physical Review X* **9**, 041029 (2019).
- [17] Kenji Nakahira. “Identification of quantum change points for hamiltonians”. *Phys. Rev. Lett.* **131**, 210804 (2023).
- [18] Gael Sentís, Emilio Bagan, John Calsamiglia, Giulio Chiribella, and Ramon Muñoz-Tapia. “Quantum Change Point”. *Physical Review Letters* **117**, 150502 (2016).
- [19] János A. Bergou, Ulrike Futschik, and Edgar Feldman. “Optimal Unambiguous Discrimination of Pure Quantum States”. *Physical Review Letters* **108**, 250502 (2012).
- [20] Gael Sentís, John Calsamiglia, and Ramon Muñoz-Tapia. “Exact Identification of a Quantum Change Point”. *Physical Review Letters* **119**, 140506 (2017).
- [21] Marco Fanizza, Michalis Skotiniotis, John Calsamiglia, Ramon Muñoz-Tapia, and Gael Sentís. “Universal algorithms for quantum data learning”. *Europhysics Letters* **140**, 28001 (2022).
- [22] Eiichi Bannai, Etsuko Bannai, Tatsuro Ito, and Rie Tanaka. “Algebraic combinatorics”. Walter de Gruyter GmbH & Co KG. (2021).
- [23] Paul Hausladen and William K. Wootters. “A ‘Pretty Good’ Measurement for Distinguishing Quantum States”. *Journal of Modern Optics* **41**, 2385–2390 (1994).
- [24] Nicola Dalla Pozza and Gianfranco Pierobon. “Optimality of square-root measurements in quantum state discrimination”. *Physical Review A* **91**, 042334 (2015).
- [25] Roelof Koekoek and René F. Swarttouw. “The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue” (1996). [arXiv:math/9602214](https://arxiv.org/abs/math/9602214).
- [26] George E. Andrews, Richard Askey, and Ranjan Roy. “Special functions”. Encyclopedia of Mathematics and its Applications. Cambridge University Press. (1999).
- [27] Stephen P Boyd and Lieven Vandenberghe. “Convex optimization”. Cambridge university press. (2004).
- [28] John Watrous. “The theory of quantum information”. Cambridge university press. (2018).
- [29] Bruce E. Sagan. “The Symmetric Group”. Volume 203 of Graduate Texts in Mathematics. Springer. New York, NY (2001).
- [30] A. S Holevo. “Statistical decision theory for quantum systems”. *Journal of Multivariate Analysis* **3**, 337–394 (1973).
- [31] Santiago Llorens, Walther González, Gael Sentís, John Calsamiglia, Emili Bagan, and Ramon Muñoz-Tapia. “Quantum edge detection” (2024). [arXiv:2405.11373](https://arxiv.org/abs/2405.11373).