Journal of Iraqi Al-Khwarizmi (JIKh) Volume: 10 Issue: 2 Year: 2026 pages: 259-280

Probabilistically Certified Randomized Krylov Recycling for Many-Query Parametric PDE Systems

Authors Names	ABSTRACT
Yahya Qader Kareem *Publication data: 26/8/2026* **Keywords:** *Parametric PDEs; Krylov subspace recycling; randomized sketching; probabilistic certification; GMRES; many-query systems.*	Many-query parametric partial differential equation workflows require repeated solutions of large sparse linear systems together with a reliable stopping statement for every query. Recent literature has developed reduced Krylov bases, randomized/sketched Krylov methods, recycling and deflation, and probabilistic model certification, but these components remain insufficiently integrated in one online solver. This paper proposes Probabilistically Certified Randomized Krylov Recycling (PCR-KR). The method warm-starts each query, constructs a local POD recycle space from recently certified solution corrections, applies a sketched coarse correction with an acceptance gate, and performs restarted sketched minimal-residual iterations. Correctness is separated from acceleration through a fresh independent Gaussian validation sketch. Under a subspace-embedding event, the sketched step is quasi-optimal and the recycle gate implies a true-residual bound. Independently, a one-sided chi-square certificate provides an upper residual bound; allocating a total failure budget over all planned queries and checks gives a familywise guarantee, and a stability lower bound converts it to a PDE state-error certificate. Fixed-seed convection–diffusion–reaction experiments cover 120 systems with 400–1600 unknowns. All queries converged with 100% observed certificate coverage. Relative to cold-start GMRES, PCR-KR reduced mean matrix–vector products by 23.36%–33.52%; it improved on warm-start GMRES in three of four principal cases. Mean residual-certificate effectivity was 1.84–1.93, with a maximum below 2.43. Deterministic recycling remained 7.53%–11.89% cheaper in matrix–vector products, quantifying the price of randomization and certification.

1. Introduction

Several large sparse linear system (LSL) problems with different coefficient matrices and vectors must be solved in series, each of which corresponds to a different physical/geometrical or control parameter in many-query PDE problems. Common applications include parameter studies, solving inverse problems, uncertainty quantification, optimization, and updating in digital twins. After spatial discretization, the computational core has the form

$$A(\mu_j)x(\mu_j) = b(\mu_j), \qquad j = 1, \dots, Q, \tag{1}$$

Conducting such a strict smooth ordering is not always possible; instead, consecutive parameters can be near each other in value. When each system is solved alone the info obtained in previous queries is lost. This inefficiency is corrected by using Krylov subspace recycling which keeps directions useful in a sequence, and by using randomized sketching to reduce the amount of work that is required to perform the residual-minimization and orthogonalization operations. Moreover, very recent randomized Householder QR results have demonstrated that stable randomized factorizations can also be used together with Krylov computation [20]. We also take reliability to be a third requirement, namely that an online solver for a multi-query context should also return not only an approximation of the state, but also a statement about a computable residual, and, if there is a known stability lower bound, a statement about a computable PDE solution error.

There has been significant progress on each of these ingredients in recent years. The reduced Krylov bases of parametric PDEs can be used to solve a family of systems in a low dimensional space [1]. The

Department of Mathematics and Statistics, Urmia University of Technology, Urmia, Iran. Corresponding author: Yahya Qader Kareem E-mail: yahyaqadeerkareem@gmail.com

randomized GMRES and sketched projection methods are rigorous dimension-reduction methods and can be decomposed using other bases that are not traditional [2]–[8] and [20]. Recycling, and deflated restarting, have been developed for evolving linear systems [2], [9]–[12], [19] and parametric systems, as well as for high-performance implementations. Simultaneously, probabilistic and certified reduced-basis techniques have provided increased level of reliability of parameter-dependent models [14]–[17]. The results here show that randomized numerical linear algebra, recycling, and certification are all/already well-established research fields. An online many-query PDE solver however is less maturely developed and needed for the intersection.

1.1 Research gap

In reviewing the literature there were found to be four streams which were at least partially separated. Addressing this issue, the parametric reduced Krylov methods and the model-reduction methods focused on the idea of generating a re-usable global approximation space, typically offline [1], [17] and [19]. Secondly, randomised Krylov methods lower the cost of residual-minimisation or basis-construction [2]–[8] but the stopping rules remain virtually unchanged from those methods used for Krylov; they are usually traditional residual-testing only, and not a self-contained familywise probabilistic certificate [2]–[8]. Third, although sequences can be deflated, harmonic Ritz vectors, POD information or block augmentation can be used to accelerate them, but do not generally provide a new one sided residual guarantee for each adaptively stopped query [2] [9]–[12] [19]. Fourth, certified reduced models yield a range of posteriori bounds, but often certify a reduced order approximation instead of a randomized online Krylov-recycling iteration [14]–[17].

Hence, in the reviewed literature from 2021 to 2026, the lack of connection between all the following features, in one online platform, is the gap found in this study:

1. sequential reuse of recent solution corrections for a many-query parametric PDE family;
2. randomized residual minimization and a randomized acceptance test for the recycle correction;
3. an independent Gaussian validation sketch that is not reused by the solver;
4. a familywise probability statement valid over all planned queries and stopping checks;
5. conversion of the certified residual to a PDE state-error bound through a computable stability lower bound; and
6. an experimental comparison against cold-start GMRES, warm-start GMRES, and deterministic recycling under smooth and scrambled query orderings.

The novelty claim is intentionally confined to this integrated combination. The paper does not claim that Krylov recycling, randomized sketching, or probabilistic error estimation is new in isolation.

1.2 Research problem

For a sequence (1), let an approximate vector z have residual

$$r_j(z) = b\left(\mu_j\right) - A\left(\mu_j\right)z. \tag{2}$$

The research problem is to develop an online method that will reuse the information that has already been certified and lower the cost associated with residual minimization using sketches and only stop when an independent upper-bound of the probability norm $\| r_j(z) \|_2$ is below some prescribed tolerance. When the recyclable space is not helpful, queries are not presented in a favourable order and the randomized solver sketch is not perfect, the method should still remain safe. The latter, however, the correctness of the final decision must still be safeguarded through an independent validation process.

1.3 Objectives and contributions

The study is made up of five objectives. O1 is to build an integrated online fixing the sequence of parametric PDE systems via the use of recent corrections, randomized residual minimization and restarting. O2 will create embedding-based quasi-optimal, and safe recycle-acceptance bound. O3 is to obtain an independent one sided residual certificate, a familywise guarantee and the corresponding PDE solution-error bound. O4 is to measure computational work by means of three open standards. O5: To investigate the robustness against mesh size, query ordering, recycle dimension, solver-sketch dimension and validation-sketch dimension.

The principle contributions are:

- a method named **Probabilistically Certified Randomized Krylov Recycling (PCR-KR)**, in which a POD recycle space is built from recent solution corrections and used as a sketched coarse correction before restarted Krylov iterations;
- a solver sketch used only for the coarse correction and sketched minimal-residual step, separated from a fresh Gaussian validation sketch used only for certification;
- a quasi-optimality theorem under a standard subspace-embedding condition, together with a true-residual bound for the recycle acceptance gate;
- a chi-square one-sided residual certificate and a union-bound construction that controls the total failure probability across all scheduled queries and checks;
- a residual-to-state-error result for coercive or inf-sup stable discretizations;
- fixed-seed experiments on 120 sequence queries for convection–diffusion–reaction systems with $n = 400$, 900, and 1600 unknowns;
- 100% observed certificate coverage in all reported sequences, mean residual-certificate effectivity between 1.84 and 1.93, and maximum effectivity below 2.43; and
- reductions in mean matrix–vector products relative to cold-start GMRES of 30.95%, 33.52%, 27.14%, and 23.36% in the four principal test cases.

The paper is organized in the following manner. Section 2 provides a critical location of the work in recent literature. In section 3 the PDE and algebra is declared. Section 4 develops PCR-KR. The theoretical guarantees are given in Section 5. The numerical protocol is described in section 6. The results for the project are reported and interpreted in Section 7. Section 8 makes a direct connection between results and objectives and deals with limitations. The paper is followed by the conclusion in section 9.

2. Recent Literature and Critical Positioning

2.1 Parametric PDEs and reduced Krylov spaces

Li, Zikatanov, and Zuo proposed and analyzed reduced Krylov basis techniques for families of parametric PDEs with a fixed-parameter preconditioner [1]. They indicate that Krylov information created at one parameter may be utilized to create low-dimensional approximation area for numerous different instances of the parameter. The method is essentially a reduced-basis approach, in which a given high-fidelity solve is reduced to a space, and the whole family is approximated in this space.

PCR-KR takes care of another kind of operating regime. It balances its recycle quantity data on the internet with the latest certified corrections (no assumption of uniform quality of one globally constructed basis).

In order to achieve an accelerated recycling for affine parametric systems, Panagiotopoulos, Desmet and Deckers suggested a decomposition of the operations conducted in the deflation space into an offline part and an online part in [19]. The present work is particularly relevant as it brings together the topics of parametric systems, model reduction and recycling. It focuses on efficient construction of a global deflation projector. The current research on the other hand adopts a smaller moving history, a randomized acceptance gate, and an independent residual certificate for each query.

McQuarrie, Khodabakhshi and Willcox created nonintrusive reduced models of parametric PDEs through operator inference [17]. Their main focus is on quick surrogate assessment following training performed on reduced operators. While it depicts the broader motivation of answering many queries, it does not capture iterative linear-system recycling, or per-query randomized residual certification.

2.2 Randomized Krylov methods

Under certain embedding conditions, ‘sketching' imposes a general framework that accelerates projection methods for linear systems and eigenvalue problems, as long as accuracy is not sacrificed [3] was proposed by Nakatsukasa and Tropp. Large-scale matrix functions were sketched approximations to FOM and GMRES by Güttel and Schweitzer [5]. A randomized version of Gram–Schmidt and subsequently a block version was developed by Balabanov and Grigori which enjoys stability and communication benefits for Arnoldi and GMRES-type methods [6, 7]. The sketch-and-select Arnoldi method was suggested by Güttel and Simunec, who used randomly added information to select a small number of vectors on which the orthogonalization is carried out [8]. Grigori and Timsit did randomized Householder QR with finite-precision and Krylov applications [20]. Bucci, Palitta, and Robol demonstrated that randomized sketched GMRES works well in linear systems of tensors, too [18].

These studies provide the numerical-linear-algebra basis of PCR-KR. Sketching is applied to the coarse problem and the residual minimisation problem and the current implementation does not support replacing Arnoldi orthogonalization with a randomized Gram–Schmidt process. It is significant as the paper does not relate communication savings to applying an orthogonalization approach which they did not actually use.

2.3 Recycling, deflation, and restarted solvers

Burke, Güttel, and Soodhalter extended the random sketching to GMRES-SDR and added deflated restarting in that method and derived some residual convergence estimates [2]. Randomised sketching was applied in Krylov recycling for matrix functions by Burke and Güttel [4]. Randomized Flexible GMRES with Deflated Restarting (RFGMRSD) was developed by Jang, Grigori, Martin and Content [9]. Thomas, Baker and Gaudreault investigated block-Arnoldi recycling with improved orthogonality [10]. Al Daas et al. [11] implemented recycling and truncation of a sequence of systems, and Bolten et al. [12] considered varying PDE structures where the structures of their discretizations may vary along a sequence.

GMRES-SDR [2] is the most recent linear-system method which bears a close relation to the present method, with the difference that it already incorporates randomisation and recycling. There are three differences between PCR-KR. First, unlike sketched harmonic Ritz vectors, the recycle space is constructed from the solution corrections that happen between the different queries. Secondly, it gets accepted or rejected via a randomized coarse-residual gate. Thirdly, only one-sided residual certificate and a familywise stopping guarantee is obtained from independent Gaussian sketch. Central is a certificate, not just the solver sketch, that serves as the reliability contribution.

2.4 Probabilistic and certified model reduction

However, a probabilistic reduced basis technique based on a PAC-type greedy construction was introduced in [14] by Billaud-Friess, Macherey, Nouy, and Prieur. The adaptive certified reduced-basis method for nonsmooth parabolic PDEs was developed by Bernreuther and Volkwein [15]. Feng et al. presented a posteriori error estimation results for parametric model reduction, and noted that there is a need for an estimator as well as computational efficiency for real applications [16]. These works are a driving force for the certification level of PCR-KR. They certify a reduced approximation computed by a reduced-basis or model-order-reduction (MOR) pipeline, while PCR-KR certifies the residual generated by an online random-Krylov (RKrylov) solve.

Randomized block adaptive solvers of linear systems have been developed by Patel, Jahangoshahi and Maldonado [13]. Their adaptive randomization corroborates the general guideline that random information might assist a solver online. This is the philosophy of PCR-KR which is based on a parametric sequence, but it also uses a separate validation sketch, thereby preventing these adaptive solver decisions from themselves constituting proof of correctness.

Table 1. Critical synthesis of the recent literature and the unresolved intersection

Ref.	Main contribution	What remains outside that study	Relation to PCR-KR
[1]	Reduced Krylov basis for parametric PDE families	No moving online recycle history or independent per-query certificate	Establishes the parametric Krylov setting
[19]	Offline/online accelerated deflation for parametric systems	Global deflation construction; no randomized validation certificate	Closest parametric recycling comparator in purpose
[2]	Randomized sketching with deflated GMRES restarting	No PDE stability conversion or familywise independent validation	Closest randomized recycling method
[3]	General sketched projection framework with accuracy guarantees	Not specialized to sequential recycling and adaptive stopping	Supplies the embedding logic used in Section 5
[4]–[8], [18], [20]	Randomized Krylov, orthogonalization, and matrix-function/tensor applications	No integrated many-query PDE certificate	Establish randomized Krylov maturity
[9]–[12]	Deflated, flexible, block, and evolving-structure recycling	Stopping usually relies on deterministic residual information	Establish recycling alternatives and reliability issues
[13]	Randomized adaptive linear-system solver	Not a Krylov-recycling framework for parametric PDEs	Motivates adaptive use of random information

Ref.	Main contribution	What remains outside that study	Relation to PCR-KR
[14]	Probabilistic reduced-basis construction	Probabilistic offline approximation, not online Krylov residual certification	Shows probability can support parametric reduction
[15], [16]	Certified reduced models and error estimators	Certification tied to ROM formulations rather than randomized recycled GMRES	Motivates explicit reliability and PDE error bounds
[17]	Data-driven parametric reduced operators	Surrogate modeling rather than sequential linear-system reuse	Reinforces the many-query application context
Present study	POD recycling, sketched solve, independent Gaussian certificate, familywise guarantee	Limited to the stated stability and experimental setting	Integrates the four streams in one online method

The table supports a precise gap statement: the contribution is not another isolated sketched GMRES variant or another reduced-basis estimator, but an integrated online path from query reuse to randomized solution and independently certified termination.

3. Problem Formulation and Assumptions

3.1 Variational and discrete problem

Let V be a Hilbert space and $\mathcal{P}$ a parameter set. For each $\mu \in \mathcal{P}$, consider: find $u(\mu) \in V$ such that

$$a_\mu(u(\mu), v) = \ell_\mu(v), \qquad v \in V. \tag{3}$$

Using a conforming or stable discrete approximation, the algebraic system of (1) is. The analysis is allowed to use nonsymmetric matrices. It is assumed that $\alpha_j > 0$ is computable, and verifies either a coercivity condition or an equivalent discrete inf-sup bound. For Euclidean coercive case,

$$z^\top \frac{A_j + A_j^\top}{2} z \geq \alpha_j \parallel z \parallel_2^2, \qquad z \in \mathbb{R}^n. \tag{4}$$

Here and below, $A_j = A(\mu_j)$ and $b_j = b(\mu_j)$.

Lemma 1 (residual-to-error stability). If (4) holds and x_j is the exact solution, then every approximation z satisfies

$$\parallel x_j - z \parallel_2 \leq \frac{\parallel r_j(z) \parallel_2}{\alpha_j}. \tag{5}$$

Proof. Let $e = x_j - z$. Since $A_j e = r_j(z)$, condition (4) and Cauchy–Schwarz give $\alpha_j \parallel e \parallel_2^2 \leq e^\top A_j e \leq \parallel e \parallel_2 \parallel r_j(z) \parallel_2$. Division by $\alpha_j \parallel e \parallel_2$ proves the claim. ▫

The same development applies when α_j is replaced by a rigorous lower bound for an inf-sup constant in compatible norms.

3.2 Many-query information model

Let $\hat{x}_i$ denote a certified approximation for query i. PCR-KR stores recent corrections

$$d_i = \hat{x}_i - \hat{x}_{i-1}. \tag{6}$$

For a history length h_q, construct

$$D_j = \left[\frac{d_{j-h_q}}{\max\left(\| d_{j-h_q} \|_2, \varepsilon_0\right)}, \dots, \frac{d_{j-1}}{\max\left(\| d_{j-1} \|_2, \varepsilon_0\right)}\right]. \tag{7}$$

The normalization prevents one unusually large parameter step from dominating the basis solely by magnitude. If

$$D_j = P_j \Sigma_j W_j^{\mathsf{T}}, \tag{8}$$

then the recycle basis consists of the first k left singular vectors whose singular values exceed a numerical rank threshold:

$$U_j = P_j(:,1{:}k_j), \qquad k_j \le k. \tag{9}$$

This construction is local in query history. It is designed for online continuation and differs from a global offline reduced basis.

3.3 Random sketches and independence

Two random maps have different responsibilities.

1. The **solver sketch** $S_j \in \mathbb{R}^{s\times n}$ is used for the recycle gate and sketched least-squares problems. In the experiments, S_j is a uniformly sampled and rescaled row operator.
2. The **validation sketch** $\Omega_j \in \mathbb{R}^{t\times n}$ is Gaussian, generated independently of S_j and of the Krylov basis construction. It is used only to certify candidate residuals.

Separating the two sketches prevents a solver from being certified by the same randomized observations that it optimized. The theoretical quality of the solver step is conditional on an embedding event for S_j; the validity of the final certificate uses only the distribution and independence of Ω_j.

4. Probabilistically Certified Randomized Krylov Recycling

4.1 Warm start and sketched recycle correction

For query $j > 1$, the initial vector is the previous certified solution,

$$x_{j,0} = \hat{x}_{j-1}, \qquad r_{j,0} = b_j - A_j x_{j,0}. \tag{10}$$

Given U_j, form $A_j U_j$ and compute the sketched coarse coefficient

$$\hat{y}_j = \operatorname{argmin}_{y\in\mathbb{R}^{k_j}} \left\| S_j\left(r_{j,0} - A_j U_j y\right)\right\|_2. \tag{11}$$

Define the observed coarse reduction ratio

$$\rho_j^{(S)} = \frac{\| S_j\left(r_{j,0} - A_j U_j \hat{y}_j\right) \|_2}{\| S_j r_{j,0} \|_2}. \tag{12}$$

The recycle correction is accepted only if $\rho_j^{(S)} \le \tau_{\text{acc}}$. Otherwise, it is discarded and the method continues from the warm start. This safeguard matters because a local history can become misleading after a large or scrambled parameter jump.

4.2 Restarted Krylov stage

After the accepted coarse correction, or directly after the warm start if it is rejected, let r be the current residual. A reorthogonalized Arnoldi process generates

$$\mathcal{K}_m(A_j, r) = \mathrm{span}\{r, A_j r, \dots, A_j^{m-1} r\}. \quad (13)$$

Let V_m be the computed basis and $W_m = A_j V_m$. PCR-KR determines

$$\hat{c} = \mathrm{argmin}_{c\in\mathbb{R}^m} \parallel S_j(r - W_m c) \parallel_2. \quad (14)$$

$x_{\text{new}} = x + V_m \hat{c}$ is the new value of the iterate. After each restart cycle true residual is re-computed to prevent any drift. As of now, sketching only reduces the dimension of the residual regressions, but does not make any "randomized-orthogonalization saving".

4.3 Independent validation and certified stopping

At every scheduled check, PCR-KR evaluates the validation sketch of the true residual. The one-sided upper estimator derived in Section 5 is denoted by $\mathcal{U}_\delta(r)$. The relative certified indicator is

$$\eta_{j,\ell}^{\text{cert}} = \frac{\mathcal{U}_{\delta_{j,\ell}}(r_{j,\ell})}{\parallel b_j \parallel_2}. \quad (15)$$

The query is accepted only when

$$\eta_{j,\ell}^{\text{cert}} \leq \varepsilon_{\text{tol}}. \quad (16)$$

The newly certified correction is then appended to the history, the oldest correction is dropped if necessary, and a new POD recycle space is generated for the next query.

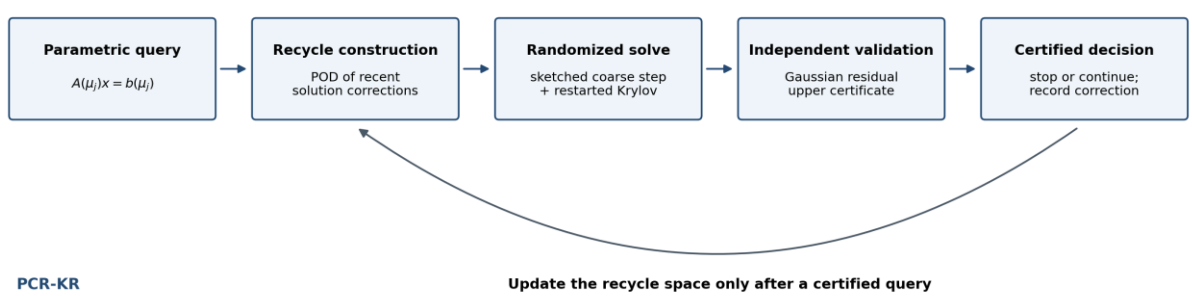

Figure 1. Workflow of PCR-KR. The solver and validation sketches have different roles, and the recycle history is updated only after a certified query.

4.4 Algorithmic statement

Algorithm 1. PCR-KR for a sequence of parametric PDE systems

Step	Operation
1	Set $x_{1,0} = 0$; for $j > 1$, set $x_{j,0} = \hat{x}_{j-1}$.
2	Build U_j by the normalized POD construction (6)–(9).
3	Generate an independent solver sketch S_j and validation sketch Ω_j.

Step	Operation
4	Solve the sketched coarse problem (11), compute (12), and accept the recycle correction only if $\rho_j^{(S)} \leq \tau_{\mathrm{acc}}$.
5	Evaluate the true residual and its independent upper certificate. Stop if (16) holds.
6	Otherwise generate an m-step Arnoldi basis, solve (14), update the iterate, recompute the true residual, and return to Step 5.
7	After certification, store $d_j = \hat{x}_j - \hat{x}_{j-1}$, truncate the history, and proceed to query $j + 1$.

The first query is solved without a recycle basis. The method is also well-defined when every recycle correction is rejected, in which case it reduces to a certified sketched restarted Krylov method with warm starts.

5. Theoretical Analysis

5.1 Subspace embedding and sketched residual minimization

A matrix $S \in \mathbb{R}^{s\times n}$ is an ε-embedding for a subspace $\mathcal{Y} \subset \mathbb{R}^n$ if

$$(1-\varepsilon) \parallel v \parallel_2^2 \leq \parallel Sv \parallel_2^2 \leq (1+\varepsilon) \parallel v \parallel_2^2, \qquad v \in \mathcal{Y}, \tag{17}$$

where $0 < \varepsilon < 1$.

Theorem 1 (quasi-optimal sketched minimal residual). Let $Z \in \mathbb{R}^{n\times q}$, and suppose S is an ε-embedding for

$$\mathcal{Y} = \mathrm{span}\{r\} + \mathrm{range}(AZ). \tag{18}$$

If

$$\hat{y} = \mathrm{argmin}_y \parallel S(r - AZy) \parallel_2, \tag{19}$$

then

$$\parallel r - AZ\hat{y} \parallel_2 \leq \sqrt{\frac{1+\varepsilon}{1-\varepsilon}} \min_y \parallel r - AZy \parallel_2. \tag{20}$$

Proof. Let y_* minimize the unsketched residual. By (17),

$\parallel r - AZ\hat{y} \parallel_2 \leq (1-\varepsilon)^{-1/2} \parallel S(r - AZ\hat{y}) \parallel_2$. Optimality of $\hat{y}$ in (19) gives an upper bound by $(1-\varepsilon)^{-1/2} \parallel S(r - AZy_*) \parallel_2$, and the upper embedding inequality gives (20). ▫

Theorem 1 is conditional: the experimental row sampler is inexpensive, but an arbitrary uniform row sample is not guaranteed to embed every possible residual subspace. The independent certificate below remains valid even if this event fails; failure affects efficiency, not the truth of the stopping bound.

Corollary 1 (benefit of exact augmentation). For any recycle space U and Krylov basis V_m,

$$\min_{y,c} \parallel r - AUy - AV_m c \parallel_2 \leq \min_c \parallel r - AV_m c \parallel_2. \tag{21}$$

This is so since the unaugmented Krylov trial space is within the augmented. It gives reasons why there is a need of establishing a deterministic recycling as an efficiency baseline. Additional work may be needed for PCR-KR as it covers the investments for randomisation and certification.

5.2 Safety of the recycle gate

Theorem 2 (true-residual implication of the sketched gate). Suppose S is an ε-embedding for $\text{span}\{r\} + \text{range}(AU)$. Let $\hat{y}$ be the sketched coarse minimizer and let $\rho^{(S)}$ be defined by (12). Then

$$\frac{\| r - AU\hat{y} \|_2}{\| r \|_2} \leq \sqrt{\frac{1+\varepsilon}{1-\varepsilon}}\, \rho^{(S)}. \tag{22}$$

Proof. The lower embedding inequality applied to the corrected residual and the upper inequality applied to r give

$\| r - AU\hat{y} \|_2 \leq \| S(r - AU\hat{y}) \|_2 / \sqrt{1-\varepsilon}$ and $\| r \|_2 \geq \| Sr \|_2 / \sqrt{1+\varepsilon}$. Their ratio proves (22). ▫

Thus an acceptance threshold τ_{acc} implies a true-residual reduction by at least the factor in (22) whenever the embedding event holds. When it does not hold, the gate is heuristic; the subsequent validation still protects termination.

5.3 One-sided Gaussian residual certificate

Let $\Omega \in \mathbb{R}^{t \times n}$ have independent entries distributed as $\mathcal{N}(0, 1/t)$. For a fixed nonzero residual r independent of Ω,

$$\frac{t \| \Omega r \|_2^2}{\| r \|_2^2} \sim \chi_t^2. \tag{23}$$

Let $\chi_{t,\delta}^2$ denote the lower δ-quantile of the chi-square distribution with t degrees of freedom.

Theorem 3 (one-sided residual upper certificate). Define

$$\mathcal{U}_\delta(r) = \sqrt{\frac{t}{\chi_{t,\delta}^2}}\, \| \Omega r \|_2. \tag{24}$$

Then

$$\Pr\{\| r \|_2 \leq \mathcal{U}_\delta(r)\} \geq 1 - \delta. \tag{25}$$

Proof. Equation (23) gives

$\Pr\{t \| \Omega r \|_2^2 / \| r \|_2^2 \geq \chi_{t,\delta}^2\} = 1 - \delta$. Rearrangement yields (25). ▫

Only the lower quantile is needed because the desired statement is a one-sided upper bound. More validation rows reduce the typical effectivity of the bound, as confirmed in Section 7.5.

5.4 Familywise validity under adaptive stopping

Visualize the entire set of candidate residuals which the solver would produce for each query without reference to the validation sketch assuming that all restart cycles were attempted. This is a “ghost sequence” which depends on A_j, b_j, the recycle history, the solver sketch and arithmetic but not on Ω_j. There is only a prefix of the actual algorithm printed because the algorithm is terminated as soon as a certificate passes.

Let $\mathcal{J}$ index all scheduled query/check pairs and choose

$$\delta_i > 0, \qquad \sum_{i \in \mathcal{J}} \delta_i \leq \delta_{\text{tot}}. \tag{26}$$

Corollary 2 (familywise certificate). With probability at least $1 - \delta_{\text{tot}}$, every scheduled candidate residual satisfies its corresponding upper bound simultaneously:

$$\Pr\{\| r_i \|_2 \leq \mathcal{U}_{\delta_i}(r_i) \text{ for all } i \in \mathcal{I}\} \geq 1 - \delta_{\text{tot}}. \tag{27}$$

Proof. Use Theorem 3 on each candidate residual and the union bound. There is no need for independence of the different events of failure. Since one of the scheduled candidates is the actual stopping residual, (27) is also true for the adaptive stopping index.

The simple allocation is $\delta_i = \frac{\delta_{tot}}{|I|}$. This conservative rule is applied to all the queries during the experiments and all viable cycle checks.

5.5 Certified PDE state error

Combining Lemma 1 and Corollary 2 gives the main PDE reliability statement.

Corollary 3 (probabilistic PDE state-error bound). Under (4), define

$$\Delta_{j,\ell}^{x} = \frac{\mathcal{U}_{\delta_{j,\ell}}(r_{j,\ell})}{\alpha_j}. \tag{28}$$

With the same familywise probability as in (27),

$$\| x_j - x_{j,\ell} \|_2 \leq \Delta_{j,\ell}^{x} \qquad \text{for every scheduled query and check.} \tag{29}$$

The residual certificate can therefore be used even when α_j is conservative. A conservative α_j may make the state-error bound loose, but it does not invalidate it.

5.6 Computational cost

Let k be the recycle dimension, h_q the history length, m the restart dimension, s the solver-sketch rows, and t the validation rows. The dominant online work per query consists of:

- an SVD of an $n \times h_q$ history matrix, $O(nh_q^2)$ for small h_q;
- k matrix–vector products to form $A_j U_j$;
- up to m matrix–vector products per restart cycle;
- classical Arnoldi orthogonalization, $O(nm^2)$ per complete cycle in the present implementation;
- a sketched least-squares solution of dimension $s \times m$, rather than $n \times m$; and
- validation products costing $O(tn)$ with the dense Gaussian sketch used here.

The matrix–vector product count is used as the primary implementation-independent work metric. Wall-clock conclusions at the tested moderate sizes would be sensitive to language, sparse-kernel implementation, and hardware and are therefore not used as the principal claim.

6. Computational Design

6.1 Parametric convection–diffusion–reaction benchmark

The benchmark is

$$-\kappa(\mu)\Delta u + \beta_x(\mu)\frac{\partial u}{\partial x} + \beta_y(\mu)\frac{\partial u}{\partial y} + \sigma u = f(\mu; x, y) \quad \text{in } \Omega = (0,1)^2, \tag{30}$$

with homogeneous Dirichlet data

$$u = 0 \quad \text{on } \partial\Omega, \tag{31}$$

The reaction coefficient is $\sigma = 0.05$. For $0 \leq \tau < 1$, the smooth query path is

$$\kappa(\tau) = 0.008 + 0.012\big(1/2 + 1/2\sin(2\pi\tau)\big). \tag{32a}$$

$$\beta_x(\tau) = 0.8 + 0.6\big(1/2 + 1/2\cos(2\pi\tau)\big). \tag{32b}$$

$$\beta_y(\tau) = 0.20 + 0.40\big(1/2 + 1/2\sin(4\pi\tau)\big). \tag{32c}$$

The moving source center is

$$x_s(\tau) = 0.35 + 0.16\cos(2\pi\tau), \qquad y_s(\tau) = 0.55 + 0.16\sin(2\pi\tau), \tag{33}$$

and

$$f(\tau; x, y) = \sin(\pi x)\sin(\pi y) + 0.35\exp\left[-90\left(\big(x - x_s(\tau)\big)^2 + \big(y - y_s(\tau)\big)^2\right)\right]. \tag{34}$$

The diffusion term uses a five-point stencil and the positive convection terms use first-order upwinding on an $N \times N$ interior grid, so $n = N^2$. If $h = 1/(N+1)$ and $\theta = \pi/(N+1)$, a lower bound for the symmetric part is

$$\alpha(\tau) = \sigma + \frac{8\kappa(\tau)}{h^2}\sin^2\left(\frac{\theta}{2}\right) + \frac{\beta_x(\tau) + \beta_y(\tau)}{h}(1 - \cos\theta). \tag{35}$$

This bound is used only to convert the residual certificate to a state-error bound. Reference solutions are computed by a sparse direct solver for validation and are not part of PCR-KR.

6.2 Methods and parameters

Four solvers are compared.

1. **GMRES(24), zero start:** restarted GMRES-like minimal-residual iteration from zero for every query.
2. **GMRES(24), warm start:** the previous solution is used as the next initial vector.
3. **Deterministic recycling:** the same warm start and POD recycle construction as PCR-KR, but exact residual least squares are used and the coarse correction is always deterministic.
4. **PCR-KR:** the proposed method with a randomized solver sketch and independent Gaussian validation.

Table 2. Fixed experimental parameters

Quantity	Value	Role
Restart dimension m	24	Maximum Arnoldi vectors per cycle
Maximum restart cycles	15	Work cap per query
Recycle dimension k	8	Maximum POD directions
History length h_q	8	Recent corrections retained
Solver-sketch rows s	160	Sketched coarse and residual regressions
Validation rows t	64	Independent Gaussian certificate
Recycle threshold τ_{acc}	0.90	Rejects weak sketched coarse corrections
Relative tolerance ε_{tol}	10^{-8}	Certified stopping target
Familywise failure budget δ_{tot}	10^{-6}	Shared by all planned checks
Random seed	12345	Reproducibility

The principal smooth sequences use $(N, Q) = (20,40)$, $(30,30)$, and $(40,20)$. A second $N = 30$, $Q = 30$ experiment applies a fixed random permutation to the same parameter points to test sensitivity to query ordering.

6.3 Metrics

For method M, the relative work reduction against baseline B is

$$\mathcal{R}(M; B) = 100\left(1 - \frac{\overline{\mathrm{MV}}(M)}{\overline{\mathrm{MV}}(B)}\right)\%, \tag{36}$$

where $\overline{\mathrm{MV}}$ is the mean matrix–vector product count per query. Certificate quality is measured by the residual effectivity

$$\eta_r = \frac{\mathcal{U}_\delta(r)}{\| r \|_2}. \tag{37}$$

Coverage corresponds to the percentage of queries for which the reported upper bounds lower and upper than the reference solution error and true residual are actually true. Since the target guarantee is probabilistic, not deterministic, the evidence of calibration is not 100% coverage, but is sharing of empirical coverage.

7. Results and Interpretation

7.1 Overall efficiency and convergence

All four methods converged on every query in all four principal cases. Table 3 gives the mean matrix–vector products.

Table 3. Mean matrix–vector products and relative PCR-KR work

Case	Cold	Warm	Det. recycle	PCR-KR	Red. vs. cold (%)	Red. vs. warm (%)	Overhead vs. det. (%)
$n = 400$, smooth, $Q = 40$	102.25	81.00	63.10	70.60	30.95	12.84	11.89
$n = 900$, smooth, $Q = 30$	196.00	138.50	118.63	130.30	33.52	5.92	9.83
$n = 900$, scrambled, $Q = 30$	196.00	165.17	132.80	142.80	27.14	13.54	7.53
$n = 1600$, smooth, $Q = 20$	241.00	182.25	169.70	184.70	23.36	-1.34	8.84

Reduction rates of 23.36-33.52% could be obtained with PCR-KR consistently better than cold-start GMRES. The method improves three cases and is almost equal at n = 1600 achieving 1.34% more matrix–vector products against warm-start GMRES. It is scientifically significant because it is an overstatement to say recycling plus certification is invariably less expensive than a ‘warm start' at each tested size. In all four cases, the deterministic recycle method has the lowest possible matrix–vector count, which is a direct consequence of the result in Corollary 1 and the fact that in this method exact residual regressions are performed without using an independent stopping certificate. PCR-KR pays an additional 7.53% - 11.89% matrix–vector work on top of this idealized deterministic baseline in return for randomly reduced dimensions and a statement of probabilistic correctness.

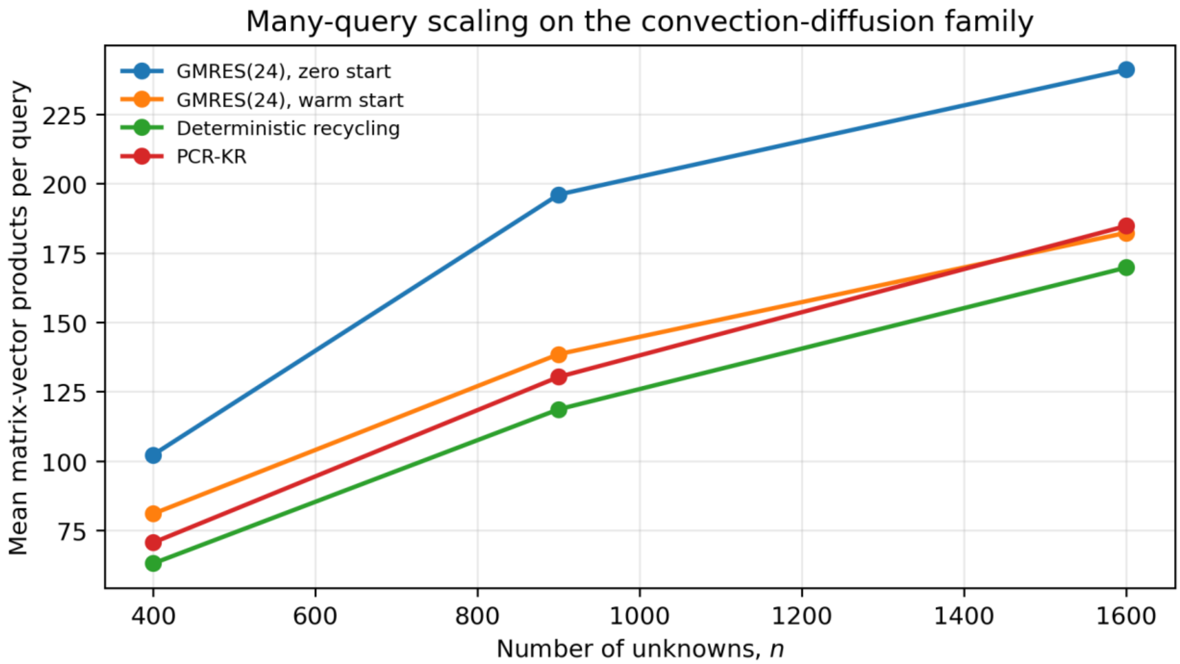

Figure 2. Mean matrix–vector products per query for the smooth parameter sequences. PCR-KR remains below cold-start GMRES at every tested size; deterministic recycling is the lowest-work baseline.

The scaling plot is consistent with the table. Finally, at n=400 and 900, PCR-KR is also slower than the warm-start GMRES. For numbers of wells n>1600 the PCR-KR and the warm-start curves are almost identical. The results suggest that the fixed parameters used (k=8, s=160, and m=24) are not generally the best for all sizes, and hence the ablation analysis in Section 7.4.

7.2 Per-query behavior

Figure 3 reports the work for each query in the n = 900 smooth sequence. The cold-start method is consistently most expensive. Warm starting removes a substantial fraction of the work, but the benefit varies with the changing diffusion, transport, and source location. Deterministic recycling remains the most economical on most queries. PCR-KR follows it closely, with occasional higher peaks caused by a rejected or less effective recycle correction and by an additional restart cycle required to satisfy the independent certificate.

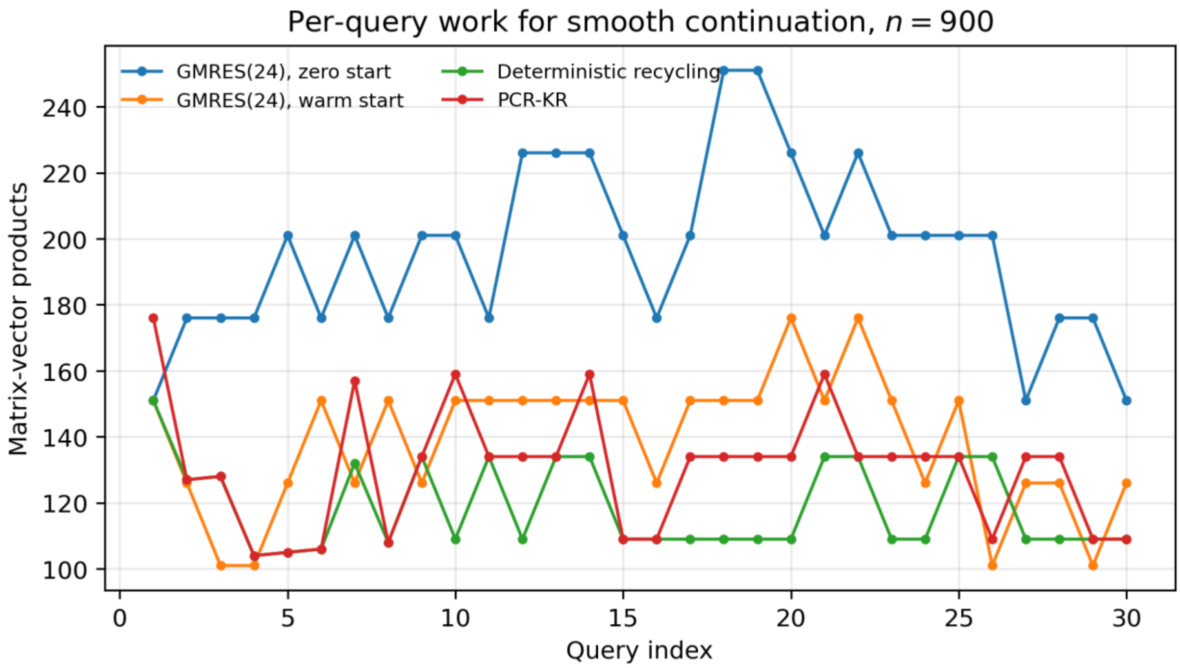

Figure 3. Per-query matrix–vector products for the smooth n = 900 sequence. The variation reflects changing parameter distance and recycle-space relevance.

The first query is not representative because it doesn't contain any history for recycling. Further queries reveal that the local POD directions provide a valuable component of the evolving solution manifold without the use of an overall reduced model. This does directly support O1 and O4.

7.3 Certificate reliability and affectivity

Table 4. PCR-KR convergence, recycle acceptance, and residual-certificate quality

Case	Queries	Converged	Accepted	Coverage	Mean η_r	Max. η_r
$n = 400$, smooth	40	40	0.9500	1.0000	1.8939	2.3997
$n = 900$, smooth	30	30	0.9333	1.0000	1.8376	2.1829
$n = 900$, scrambled	30	30	0.9667	1.0000	1.9294	2.2424
$n = 1600$, smooth	20	20	0.9000	1.0000	1.9097	2.4225

For each of 120 sequence queries, all the certified residual upper bounds contained the true residual, and all the state-error bounds contained the reference error. The average effectivity of the remains is between 1.84 and 1.93 and the greatest recorded value is 2.4225. Though these values are lower bounds, they are nonetheless useful; the certificate is not simply a useful order-of-magnitude gauge.

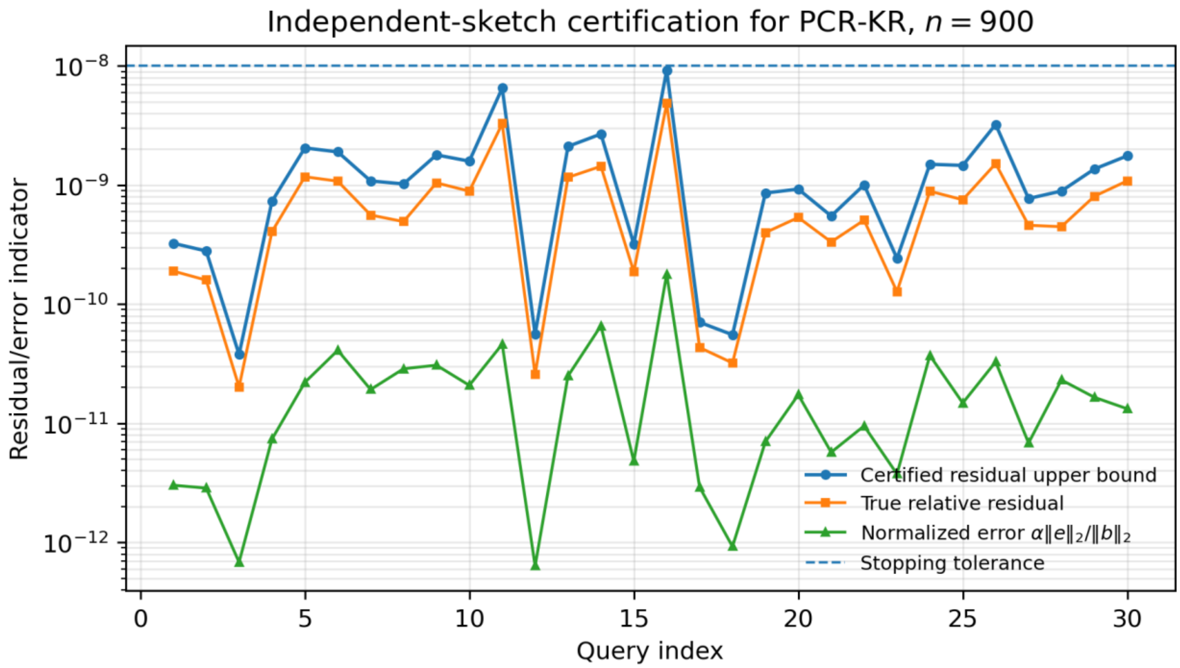

Figure 4. Independent residual certification for the smooth n = 900 sequence. The certified upper bound remains above the true residual, and the stability-scaled true state error remains below both.

Figure 4 distinguishes three quantities. The orange curve is the true relative residual, available only for validation. The blue curve is the independently certified upper bound used by the algorithm. The green curve is $\alpha_j \| e_j \|_2 / \| b_j \|_2$; Lemma 1 predicts that it cannot exceed the true relative residual. The observed ordering supports O3 and verifies the direct link between the probability theorem and the PDE stability objective.

7.4 Sensitivity to query ordering and hyper parameters

When you start the computation for n=900 queries, the mean number of warm-start work is 165.17 matrix–vector products, which is higher than the previous 138.50 matrix–vector products when n=100 queries, because two close solves are no longer following the smooth parameter path. The PCR-KR goes up from 130.30 to 142.80, a position that still means it's 13.54% better than the scrambled start method run with a warm-start. Its recycle acceptance rate increases a bit, but there is no such thing as a quality assurance from its acceptance only; an accepted subspace may need additional Krylov work after the coarse stage.

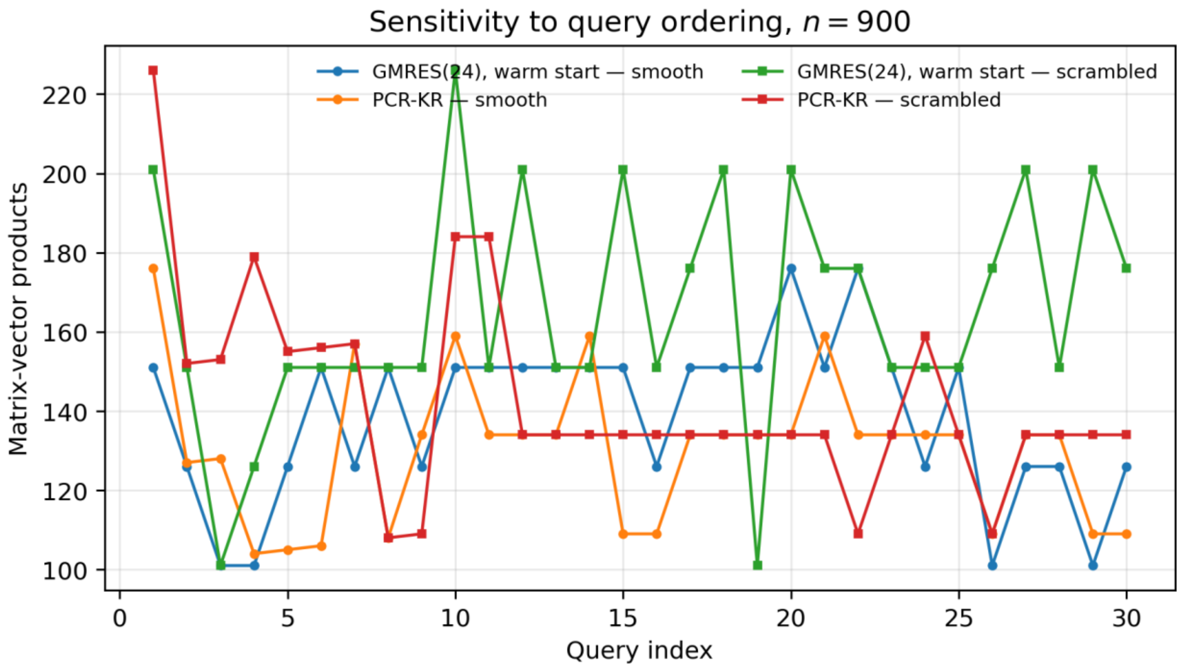

Figure 5. Sensitivity to query ordering at n = 900. Scrambling weakens simple warm starts more strongly than PCR-KR.

Table 5 reports a focused ablation at $N = 30$, $Q = 20$ on the smooth sequence.

Table 5. Ablation of recycle dimension and solver-sketch rows

Parameter	Value	Mean matrix–vector products	Acceptance	Coverage	Mean η_r	Maximum η_r
Recycle dimension	0	149.75	0.00	1.00	1.8196	2.2442
Recycle dimension	4	147.00	0.90	1.00	1.8497	2.1965
Recycle dimension	8	140.95	0.90	1.00	1.8347	2.0239
Recycle dimension	12	142.85	0.90	1.00	1.8306	2.0513
Solver-sketch rows	64	144.70	0.90	1.00	1.8512	2.3114
Solver-sketch rows	96	142.20	0.90	1.00	1.8718	2.0734
Solver-sketch rows	160	140.95	0.90	1.00	1.8347	2.0239
Solver-sketch rows	256	135.95	0.90	1.00	1.8530	2.2234

The optimum value of k in the recycle study is nonmonotone with respect to the values tested, with k=8 yielding a better result than k=4 and k=12. Having a larger recycle space is not an advantage as it means you will have to pay extra products $A_j U_j$, it may even involve stale directions and the coarse regression dimension is larger. As s increases, matrix – vector work tend to go down, sketch processing goes up. Thus based on Table 5, one can suggest the selection of the adaptive parameters as further work instead of mentioning that s=256 is optimum in all cases.

7.5 Monte Carlo calibration of the one-sided certificate

There is an independent experiment involving 50,000 trials that separates Theorem 3. The nominal coverage was 0.99 for fixed residual norm, and for independent Gaussian sketches.

Table 6. Monte Carlo calibration of the one-sided Gaussian upper certificate

Validation rows t	Trials	Nominal coverage	Observed coverage	Median upper effectivity	95th percentile	Maximum
24	50,000	0.99	0.99034	1.4673	1.8301	2.4154
48	50,000	0.99	0.99032	1.2957	1.5200	1.8974
64	50,000	0.99	0.99054	1.2471	1.4342	1.7365
96	50,000	0.99	0.99040	1.1957	1.3397	1.5762
128	50,000	0.99	0.98964	1.1661	1.2884	1.4575

Observed coverage varies only within sampling variation of 0.99 at each dimension. The effectivities at the median and upper tail both decrease with increasing t, showing the clarity of performance and reliability–performance conflict.

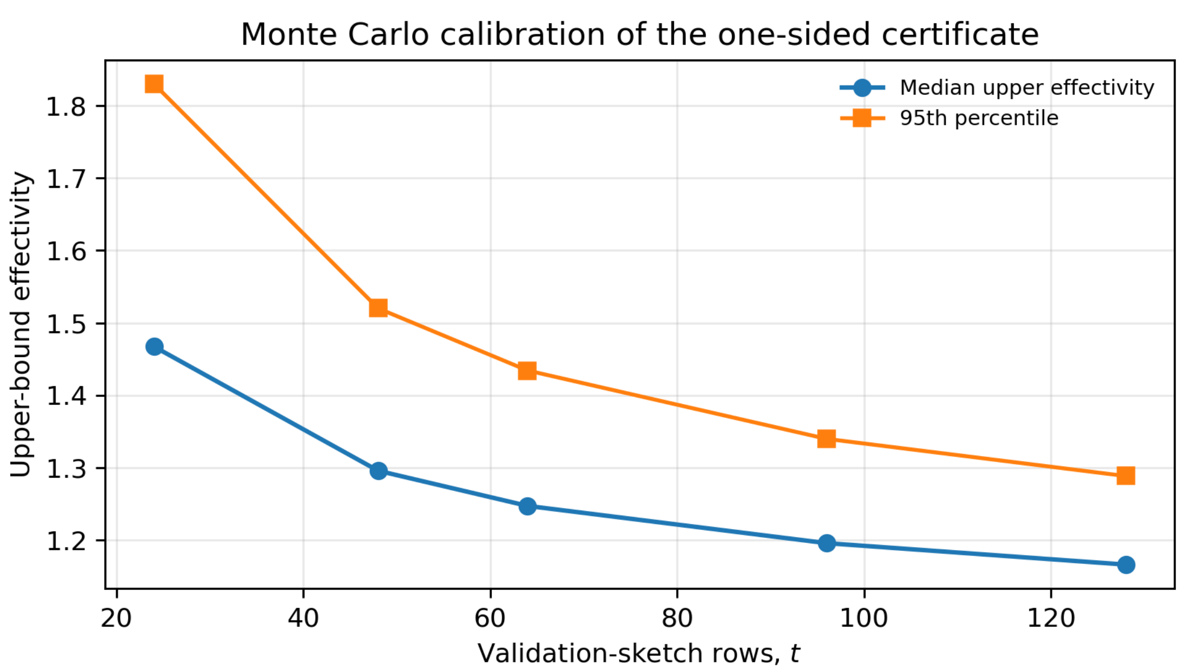

Figure 6. Calibration of the one-sided Gaussian certificate. Additional validation rows tighten the bound without changing its nominal probability statement.

The sequence experiments are run with t=64 and on a much smaller per-check failure probability that 0.01 (since the entire budget is spread out over all the planned checks, 10^{-6}. The greater sequence-level residual effectivities, which are evident in Table 6 compared with the median values, are an explanation of this conservative familywise allocation.

7.6 Direct objective–result alignment

Table 7. Direct connection between objectives, evidence, and interpretation

Objective	Evidence in the study	Interpretation
O1: integrated online solver	Algorithm 1; Figures 1 and 3	Recent corrections, sketching, restarting, and certification operate in one sequential workflow.
O2: solver and recycle guarantees	Theorems 1–2 and Corollary 1	Sketched residual minimization is quasi-optimal on an embedded trial space; the gate has a true-residual implication.
O3: independent familywise certificate	Theorem 3, Corollaries 2–3, Table 4, Figure 4	The stopping decision has a probability budget independent of the solver sketch and converts to a PDE error statement.
O4: efficiency against baselines	Table 3 and Figure 2	PCR-KR reduces work relative to cold GMRES in all cases and relative to warm GMRES in three cases, while deterministic recycling remains cheaper.
O5: robustness and sensitivity	Figure 5; Tables 5–6	The method tolerates scrambled ordering; recycle and sketch dimensions affect efficiency; validation rows control effectivity.

This organization prevents the numerical section from becoming a detached collection of performance plots. Every principal finding answers an objective stated before the method was introduced.

8. Discussion

8.1 Relationship to previous studies

The findings are in accordance with the key-objective of the reduced Krylov basis study [1] that there is information of low dimension behind the parametric PDE solution families that can be reused. PCR-KR adopts this information locally and not a single offline "basis" around the world. The nonmonotone result on the recycle dimension in Table 5 also renders the following observation: When query paths change, or are scrambled, then a fixed global space is not always an optimal choice when operating online.

Compared with the parametric recycling strategy in [19] PCR-KR eliminates the requirement of an affine offline/online projector decomposition and updates a small POD basis directly from latest corrections. While this is no longer as applicable as the other one to the nonaffine moving-source right-hand side of (34), it foregoes the large offline amortization that comes when a global affine model is accurate.

The method proposed in the present paper is simpler than GMRES-SDR [2] (also called Single Domain Restricted Arnoldi in the next text) and randomized flexible GMRES [9] (also called Single Domain Randomized Flexible Arnoldi in the next text) with respect to the construction of the recycle and with respect to the use of a conventional Arnoldi stage. It's not about the deflation method being stronger: We have shown that for cases involving matrix-vector products, deterministic recycling is clearly less expensive. Its distinguishing features are independent validation sketch and the familywise link to PDE state error. In Table 3 it can be observed that PCR-KR does not claim to outperform the best deterministic recycler, but rather adds a clearly stated and measured reliability layer.

The embedding analysis itself is a sketched projection viewpoint of [3] and works with randomized Gram–Schmidt or sketch-and-select Arnoldi [6]–[8]. Those basis-construction techniques would decrease the amount of synchronization that is required in a parallel implementation, but is not accounted for in the present experiments. The tensor of [18] further suggests that the solver-sketch element may be more general than sparse matrices.

Both the probabilistic reduced-basis approach [14] and certified approaches [15] and [16] show that reliability and probabilistic approximation is possible in parameter-dependent simulation. PCR-KR is a new paradigm that relaxes this principle to online Krylov termination. A minimised, random, solver residual, calculated using the sketch norm, should not be assumed to be a random and unbiased certificate of the full residual of the same sketch.

8.2 Strengths

First Strength is the Logical Separate. The independence of the random information is used to both validate (check) correctness in the solver's answers and for the Solver acceleration. However, a poor solver sketch may lead to more work, or to a weaker coarse correction which may be necessary; it does not however make the stopping statement invalid. Direct familywise control is a second strength. The total risk is assigned prior to sequence and adaptive stopping is over a ghost sequence of candidate residuals. The third strength is the transparent PDE interpretation done in Lemma 1 of the article and Corollary 3. The fourth strength is that honest benchmarking (cold-start, warm-start and deterministic recycling) considers the value of continuation, recycling and certification separately. The fifth is about reproducibility: with fixed seeds and path of parameters to be followed.

8.3 Limitations and threats to validity

Several limitations delimit the conclusions.

First is the maximum size reported is 1600, which is enough to test for algorithmic mechanisms but is not a production size HPC sized PDE solve system. Therefore the focus is not so much wall-clock speedup, but rather matrix–vector products. Second, there is no preconditioner! Preconditioner reuse will be a key step forward and can be combined with the certificate by industrial PDE solvers, which typically involve parameter-robust preconditioning. Third, uniform row sampling is an experiment to perform as a solver sketch. The final sketch is a subspace embedding, but not at each iteration as indicated in Theorem 1. The independence of the validation certificate guarantees correctness and not quasi-optimal efficiency. Fourth, Dense Gaussian validation product has O(tn) per check costs. This could be reduced using structured Gaussian compatible or fast Johnson–Lindenstrauss transforms, but these would need to be analysed separately as a certificate model. For example, in the fifth experiment we use one family of convection–diffusion–reaction equations. Experimental testing of other PDE classes, strongly indefinite operators and change of meshes. Sixth, the error in between the state and the reference can be very high if the stability bound (35) is loose. Finally, in each principal case, merely fewer matrix–vector products are needed for deterministic recycling. In choosing between PCR-KR and ordinary counting, keep in mind using PCR-KR: An explicit probabilistic stopping statement may be more important than the number of iterations alone.

The last validity point is novelty. The literature stream is constantly changing. Gap statement presented in the paper should be rechecked before submission against the published work in Scopus, Web of

Science, MathSciNet and against newly published in the field that is specially reviewed in this paper.

9. Conclusions

In this study, an online solver, called PCR-KR, was developed for many-query parametric PDE systems based on warm starts, a POD recycle space identified from the recent problem history with respect to the problem solutions, randomized residual minimization, with a recycle-acceptance gate, and an independent Gaussian validation certificate. In a subspace-embedding event, the sketched residual step is effectively 'quasi optimal' and the gate gives rise to an upper bound on the actual coarse residual. The Gaussian validation sketch, independently of that event, provides a one sided residual upper bound. If a total failure budget is allocated among all scheduled queries and checks, a familywise guarantee can be maintained until the adaptive stopping index. A stability lower bound is then used to only transform the remainder certificate into a PDE state error certificate.

All 120 sequence queries converged in the fixed-seed convection–diffusion–reaction experiments and 100% certificate coverage was observed. When using PCR-KR, the average reductions of GMRES with cold-start is 23.36%-33.52% for matrix–vector work. It has also achieved some further enhancement over warm-start GMRES in three out of four main cases; in the fourth case, it was within 1.34%, also for n=1600. Deterministic recycling saved 7.53%–11.89% on the number of matrix–vector products compared to PCR-KR, but only in a quantifiable way. Mean residual effectivity was in the range from 1.84 to 1.93 and the calibration on a separate trial of 50,000 agreed with the nominal coverage of the one-sided chi-square certificate.

The results address the research problem in the tested regime: recycling and randomized Krylov computation can be linked to an independent (familywise) PDE-interpretable certification mechanism. Future research should look at other options for adaptive risk allocation, leverage-aware or structured solver sketches, randomized low-synchronization Arnoldi, parameter-robust preconditioner recycling, changing meshes, block right hand sides, and large-scale parallel benchmarks.

References

[1] Y. Li, L. T. Zikatanov, and C. Zuo, “Reduced Krylov basis methods for parametric partial differential equations,” *SIAM J. Numer. Anal.*, vol. 63, no. 2, pp. 976–999, 2025, doi: 10.1137/24M1661236.

[2] L. Burke, S. Güttel, and K. M. Soodhalter, “GMRES with randomized sketching and deflated restarting,” *SIAM J. Matrix Anal. Appl.*, vol. 46, no. 1, pp. 702–725, 2025, doi: 10.1137/23M1619472.

[3] Y. Nakatsukasa and J. A. Tropp, “Fast and accurate randomized algorithms for linear systems and eigenvalue problems,” *SIAM J. Matrix Anal. Appl.*, vol. 45, no. 2, pp. 1183–1214, 2024, doi: 10.1137/23M1565413.

[4] L. Burke and S. Güttel, “Krylov subspace recycling with randomized sketching for matrix functions,” *SIAM J. Matrix Anal. Appl.*, vol. 45, no. 4, pp. 2243–2262, 2024, doi: 10.1137/23M1595904.

[5] S. Güttel and M. Schweitzer, “Randomized sketching for Krylov approximations of large-scale matrix functions,” *SIAM J. Matrix Anal. Appl.*, vol. 44, no. 3, pp. 1073–1095, 2023, doi: 10.1137/22M1518062.

[6] O. Balabanov and L. Grigori, “Randomized Gram–Schmidt process with application to GMRES,” *SIAM J. Sci. Comput.*, vol. 44, no. 3, pp. A1450–A1474, 2022, doi: 10.1137/20M138870X.

[7] O. Balabanov and L. Grigori, “Randomized block Gram–Schmidt process for the solution of linear systems and eigenvalue problems,” *SIAM J. Sci. Comput.*, vol. 47, no. 1, pp. A553–A585, 2025, doi: 10.1137/22M1518396.

[8] S. Güttel and I. Simunec, “A sketch-and-select Arnoldi process,” *SIAM J. Sci. Comput.*, vol. 46, no. 4, pp. A2774–A2797, 2024, doi: 10.1137/23M1588007.

[9] Y. Jang, L. Grigori, E. Martin, and C. Content, “Randomized flexible GMRES with deflated restarting,” *Numer. Algorithms*, vol. 98, no. 1, pp. 431–465, 2025, doi: 10.1007/s11075-024-01801-3.

[10] S. J. Thomas, A. H. Baker, and S. Gaudreault, “Augmented MGS-CGS block-Arnoldi recycling solvers,” *SIAM J. Sci. Comput.*, vol. 47, no. 3, pp. A1458–A1485, 2025, doi: 10.1137/23M1598544.

[11] H. Al Daas, L. Grigori, P. Hénon, and P. Ricoux, “Recycling Krylov subspaces and truncating deflation

subspaces for solving sequence of linear systems," *ACM Trans. Math. Softw.*, vol. 47, no. 2, pp. 1–30, 2021, doi: 10.1145/3439746.

[12] M. Bolten, E. de Sturler, C. Hahn, and M. L. Parks, "Krylov subspace recycling for evolving structures," *Comput. Methods Appl. Mech. Eng.*, vol. 391, Art. no. 114222, 2022, doi: 10.1016/j.cma.2021.114222.

[13] V. Patel, M. Jahangoshahi, and D. A. Maldonado, "Randomized block adaptive linear system solvers," *SIAM J. Matrix Anal. Appl.*, vol. 44, no. 3, pp. 1349–1369, 2023, doi: 10.1137/22M1488715.

[14] M. Billaud-Friess, A. Macherey, A. Nouy, and C. Prieur, "A probabilistic reduced basis method for parameter-dependent problems," *Adv. Comput. Math.*, vol. 50, Art. no. 19, 2024, doi: 10.1007/s10444-024-10114-x.

[15] M. Bernreuther and S. Volkwein, "An adaptive certified space-time reduced basis method for nonsmooth parabolic partial differential equations," *Adv. Comput. Math.*, vol. 50, Art. no. 48, 2024, doi: 10.1007/s10444-024-10137-4.

[16] L. Feng, S. Chellappa, and P. Benner, "A posteriori error estimation for model order reduction of parametric systems," *Adv. Model. Simul. Eng. Sci.*, vol. 11, Art. no. 5, 2024, doi: 10.1186/s40323-024-00260-8.

[17] S. A. McQuarrie, P. Khodabakhshi, and K. E. Willcox, "Nonintrusive reduced-order models for parametric partial differential equations via data-driven operator inference," *SIAM J. Sci. Comput.*, vol. 45, no. 4, pp. A1917–A1946, 2023, doi: 10.1137/21M1452810.

[18] A. Bucci, D. Palitta, and L. Robol, "Randomized sketched TT-GMRES for linear systems with tensor structure," *SIAM J. Sci. Comput.*, vol. 47, no. 5, pp. A2801–A2827, 2025, doi: 10.1137/24M1694999.

[19] D. Panagiotopoulos, W. Desmet, and E. Deckers, "An accelerated subspaces recycling strategy for the deflation of parametric linear systems based on model order reduction," *Comput. Methods Appl. Mech. Eng.*, vol. 403, Art. no. 115765, 2023, doi: 10.1016/j.cma.2022.115765.

[20] L. Grigori and E. Timsit, "Randomized Householder QR," *SIAM J. Sci. Comput.*, vol. 48, no. 3, pp. A1611–A1642, 2026, doi: 10.1137/24M1674327.