Pseudo general parameter shift rule

Vu Tuan Hai
University of Information Technology, Ho Chi Minh City, 700000, Vietnam
Vietnam Nation University, Ho Chi Minh City, 700000, Vietnam
haivt@uit.edu.vn
Phan Hoang Chuong
University of Information Technology, Ho Chi Minh City, 700000, Vietnam
Vietnam Nation University, Ho Chi Minh City, 700000, Vietnam
chuongph@uit.edu.vn
Pham The Bao
Saigon University, Ho Chi Minh City, 700000, Vietnam
ptbao@sgd.edu.vn

***Abstract*—The parameter-shift rule is an approach to evaluating gradients in many variational quantum algorithms. Although the general parameter-shift rule has been proposed, it needs at least linear quantum evaluations regarding the number of eigenvalues of the generator of the quantum gate. So, here we discuss an approximate method named the pseudo-parameter-shift rule to reduce the number of quantum evaluations.**

***Index Terms*—quantum computing, parameter-shift rule**

I. Introduction

Variational quantum algorithms [1], [2] — have come to the popular tool in the noisy intermediate-scale quantum (NISQ) era. They have used to solve various problems from combination optimization [8], compilation [15], quantum control [16] to quantum machine learning [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. The idea behind VQA is just simply and follow a strategy: an ansatz (or parameterized quantum circuit - PQC) is initialized. When it's executed by many measurements, we get the value of cost function as expected value. After that, some quantum / classical optimization is used to minimize that cost function. Initially, this task is treated as black-box optimization and gradient-free optimizers such as Nelder-Mead or COBYLA, were applied. However, like machine learning [33, 34, 35, 36, 37], the first-order gradient optimizers include: stochastic gradient descent (SGD), Adaptive Moment Estimation (Adam), ... provides significant advantages, via a technique named parameter-shift rule (PSR) [16, 23, 38, 39].

Although gradient-based methods are better but it also require more computation resources because the PSR need some evaluation on the PQC. The PSR formula is originally used only the case of quantum gate has two distinct eigenvalues and just developed for R number of distinct eigenvalues by discrete Fourier transform (DFT), so-called general PSR. The general PSR require at least $\mathcal{O}(2R+1)$ quantum evaluation with increase linearly wrt the size of quantum gate. In this paper, we proposed a approximate method that reduce the number of evaluation to constant.

II. Background

A. VQE problem

Consider a variational circuit $U(\boldsymbol{\theta})$ parametrized by $\theta_0, \theta_1, \theta_2, ...$ It can be contains a sequence of fixed and parametrized layer, called n is the number of qubit. We can rewrite $U(\boldsymbol{\theta})$ as:

$$U(\theta) = U_L(\boldsymbol{\theta}_L)V_L U_{L-1}(\boldsymbol{\theta}_{L-1})V_{L-1}...U_0(\boldsymbol{\theta}_0)V_0 \quad (1)$$

A fixed gate can be treated as identity operator, so it does not affect to computing gradient process. All $U_\ell(\boldsymbol{\theta}_\ell) = \{U_j\}_{j=1}^n$ is the combination of the below common parameterized gates.

$$\{R_X, R_Y, R_Z, CR_X, CR_Y, CR_Z, R_{XX}, R_{YY}, R_{ZZ}\}$$

where each gates can be represented at form $e^{-i\theta G}$ with G as the generator, such that $U_j(\theta) = e^{-i\theta G} = \cos(\theta)\mathbb{I} - i\sin(\theta)G$.

Consider the cost function of $U(\boldsymbol{\theta})$ (or U in short-handed from).

$$E(\boldsymbol{\theta}) := \langle\psi|U^\dagger \hat{B} U|\psi\rangle, \quad (2)$$

where $\hat{B}$ is the measurement operator and ψ is the initial quantum state (simply as $|0\rangle^{\otimes n}$). The problem need to solved here is finding θ such as $E(\boldsymbol{\theta})$ achieve min value:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} E(\boldsymbol{\theta}) \quad (3)$$

This task can be done iteratively by classical optimizer, simply gradient descent:

$$\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^t - \alpha \nabla_{\boldsymbol{\theta}} E(\boldsymbol{\theta}^t) \quad (4)$$

where $\alpha \in \mathbb{R}_{\geqslant 0}$ is the fixed learning rate and $\nabla E(\theta^t) = \langle\psi|\nabla(U(\boldsymbol{\theta})^\dagger \hat{B} U(\boldsymbol{\theta}))|\psi\rangle = [\partial_{\theta_0} E(\boldsymbol{\theta}^t) \;\; \partial_{\theta_1} E(\boldsymbol{\theta}^t) \;\; ...]^T$. Or Adam optimizer:

$$\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (5)$$

where $m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla_{\boldsymbol{\theta}}\mathcal{C}(\boldsymbol{\theta}), v_t = \beta_2 v_{t-1} + (1-\beta_2)\nabla^2_{\boldsymbol{\theta}}\mathcal{C}(\boldsymbol{\theta}), \hat{m}_t = m_t/(1-\beta_1^t), \hat{v}_t = v_t/(1-\beta_2^t)$, with the hyper-parameters $\beta_1 = 0.8$ or $0.9, \beta_2 \approx 1$ and $\epsilon \approx 0$.

This gradient has the exact same form as the single-gate case. In terms of the circuit, this means we can leave all other gates as they are, and only modify gate $U(\theta_i)$ when we want to differentiate with respect to the parameter θ_i

B. Parameter-shift rule

Taking first-order partial derivative of $E(\theta)$ w.r.t θ_j:

$$\begin{aligned}\partial_{\theta_j} E(\theta) &= \partial_{\theta_j}\langle\psi|U^\dagger \hat{B} U|\psi\rangle \\ &= \langle\psi|(\partial_{\theta_j}U)^\dagger \hat{B} U|\psi\rangle + \langle\psi|U^\dagger \hat{B}(\partial_{\theta_j}U)|\psi\rangle \\ &= \langle\psi|(-iGU)^\dagger \hat{B} U|\psi\rangle + \langle\psi|U^\dagger \hat{B}(-iGU)|\psi\rangle \\ &= -i\langle\psi|(U^\dagger[\hat{B}, G]U|\psi\rangle\end{aligned} \tag{6}$$

where $[\hat{B}, G] = \hat{B}G - G\hat{B}$ is the commutator. For a generator that has two distinguish eigenvalues, such as the standard rotation gates based on Pauli matrices, e.g. $G \in \{\sigma_X, \sigma_Y, \sigma_Z\}$, the equation. 6 obey to two-term parameter-shift rule:

$$\partial_{\theta_j} E(\theta) = \frac{1}{2\sin(s)}[E(\theta + se_j) - E(\theta - se_j)] \tag{7}$$

where e_j is the j^{th} unit vector and s is an freely shift, simply from 0 to π. Popularly, if $s = \frac{\pi}{2}$, we obtain the parameter-shift rule already used in libraries [22–26].

Because equation. 7 is the analytic gradient, not approximate gradient, so we can apply it again to get higher-order derivative, such as second-order derivative w.r.t to any parameters $\theta_{j_1}, \theta_{j_2}$.

$$\begin{aligned}\partial_{\theta_j\theta_k} E(\theta) &= \frac{1}{2\sin(s_1)}[\partial_{\theta_k}E(\theta + s_1 e_{j_1}) - \partial_{\theta_k}E(\theta - s_1 e_{j_1})] \\ &= \frac{1}{4\sin(s_1)\sin(s_2)}[E(\theta + s_1 e_{j_1} + s_2 e_{j_2}) \\ &\quad - E(\theta + s_1 e_{j_1} - s_2 e_{j_2}) - E(\theta - s_1 e_{j_1} + s_2 e_{j_2}) \\ &\quad + E(\theta - s_1 e_{j_1} - s_2 e_{j_2})]\end{aligned} \tag{8}$$

which is similar to the gradient formula used in Refs. [10–13], but extended to the Hessian. A formula equivalent to Eq. (11) for the particular case $s = \frac{\pi}{2}$ was recently used in Refs. [17,27].

For arbitrary-order derivative form:

$$\partial_{\theta_{j_0}\theta_{j_1}\ldots\theta_{j_n}} E(\theta) \tag{9}$$

in case any $s_j = \frac{\pi}{2}$, the equation. 9 is simply as:

$$\partial_{\theta_0\theta_1\ldots\theta_n} E(\theta) = \frac{1}{2^n}\Sigma_{k\in S_{j_1,j_2,\ldots,j_n}} P(\boldsymbol{k})E(\theta + \boldsymbol{k}) \tag{10}$$

where $\mathbf{k}_{\pm j_1,\pm j_2,\ldots,\pm j_d} = \frac{\pi}{2}(\pm\mathbf{e}_{j_1} \pm \mathbf{e}_{j_2} \pm \cdots \pm \mathbf{e}_{j_d})$, $S_{j_1,j_2,\ldots,j_n} = \{\boldsymbol{k}_{\pm j_1,\pm j_2,\ldots,\pm j_n} \; \forall$ choice of signs$\}$ and $P(\boldsymbol{k})$ is the parity of a shift vector [[**?**]].

The question here is how many expectation values are necessary to evaluate the n-order derivative $\partial_{\theta_0\theta_1\ldots\theta_n}E(\theta)$. Follow the analysis in [**?**], that is the number of elements in the sum, in equation. 10, which is $|S_{j_1,j_2,\ldots,j_n}| = 2^n$

Taking into account that the derivative tensor is symmetric with respect to permutations of the indices, the number of distinct elements of a tensor of order n is given by the combinations of n indices sampled with replacement from the set of m variational parameters. Therefore, the total number of expectation values to evaluate a derivative tensor of order n is bounded by

$$\text{No. of expectation values} \leqslant 2^n \binom{m+n-1}{n} = \mathcal{O}(m^n) \tag{11}$$

where in the last step we assumed $m \gg n$. In the opposite regime $n \gg m$, each angle parameter θ_j can be subject to multiple shifts.

This concludes the first part of this work, in which we derived several exact analytical results. However, if we really want to apply Eqs. (9), (11), and (21) in a quantum experiment, we have to take into account that each expectation value on the right-hand side of these equations can be measured only up to a finite precision and so the derivative on the left-hand side can be estimated only up to a finite error.

C. General parameter-shift rule

Let analysis the U operator again. The eigenvalues of $U(\theta)$ are given by $\{\exp(i\omega_j\theta)\}_{j\in[d]}$ with real-valued $\{\omega_j\}_{j\in[d]}$ where we denote $[d] := \{1,\ldots,d\}$ and have sorted the ω_j to be non-decreasing. Thus, we have:

$$\begin{aligned}E(\theta) &= \sum_{j,k=1}^{d} \overline{\psi_j e^{i\omega_j\theta}} b_{jk}\psi_k e^{i\omega_k\theta} \\ &= \sum_{\substack{j,k=1 \\ j<k}}^{d} \Big[\overline{\psi_j} b_{jk}\psi_k e^{i(\omega_k-\omega_j)\theta} \\ &\quad + \psi_j \overline{b_{jk}\psi_k e^{i(\omega_k-\omega_j)\theta}}\Big] + \sum_{j=1}^{d} |\psi_j|^2 b_{jj},\end{aligned} \tag{12}$$

where we have expanded B and $|\psi\rangle$ in the eigen basis of U, denoted by b_{jk} and ψ_j, respectively.

We can collect the x-independent part into coefficients $c_{jk} := \overline{\psi_j} b_{jk}\psi_k$ and introduce the R unique positive differences $\{\Omega_\ell\}_{\ell\in[R]} := \{\omega_k - \omega_j \mid j,k \in [d], \omega_k > \omega_j\}$. Note that the differences are not necessarily equidistant, and that for $r = \left|\{\omega_j\}_{j\in[d]}\right|$ unique eigenvalues of the gate generator, there are at most $R \leqslant \frac{r(r-1)}{2}$ unique differences. However, many quantum gates will yield $R \leqslant r$ equidistant differences instead; a common example for this is

$$G = \sum_{k=1}^{\mathcal{P}} \pm P_k$$

for commuting Pauli words P_k $(P_k P_{k'} = P_{k'} P_k)$, which yields the frequencies $[\mathcal{P}]$ and thus $R = \mathcal{P}$. In the following, we implicitly assume a mapping between the two indices $j, k \in [d]$ and the frequency index $\ell \in [R]$ such that $c_\ell = c_{\ell(j,k)}$ is well-defined [2]. We can then write the expectation value as a trigonometric polynomial (a finite-term Fourier series):

By using the finite-term Fourier series, we can write (2) as:

$$\begin{aligned} E(x) &= a_0 + \sum_{\ell=1}^{R} c_\ell e^{i\Omega_\ell x} + \sum_{\ell=1}^{R} \overline{c_\ell} e^{-i\Omega_\ell x} \\ &= a_0 + \sum_{\ell=1}^{R} a_\ell \cos(\Omega_\ell x) + b_\ell \sin(\Omega_\ell x) \end{aligned} \tag{13}$$

with frequencies given by the differences $\{\Omega_\ell\}_{\ell \in [R]} := \{\lambda_k - \lambda_j | j, k \in [n], \lambda_k > \lambda_j\}$ and $a_\ell, b_\ell \in \mathbb{R}$. In general, the form of $E(x)$ is known exactly as (13) can be separated as:

$$\begin{aligned} E_{\text{odd}}(x) &= \frac{1}{2}(E(x) - E(-x)) \\ &= \sum_{\ell=1}^{R} b_\ell \sin(\ell x) \end{aligned} \tag{14}$$

$$\begin{aligned} E_{\text{even}}(x) &= \frac{1}{2}(E(x) + E(-x)) \\ &= a_0 + \sum_{\ell=1}^{R} a_\ell \cos(\ell x) \end{aligned} \tag{15}$$

Since $E(x)$ is a finite-term Fourier series, the coefficients $a_\ell, b_\ell \in \mathbb{R}$ can be obtained by discrete Fourier transform (DFT). In detail, determining $E_{\text{even}(x)}$ requires $R + 1$ evaluations for coefficients $\{a_\ell\}_{\ell=0}^{R}$. Determining $E_{\text{odd}(x)}$ requires R evaluations for coefficients $\{b_\ell\}_{\ell=1}^{R}$. The total is $2R + 1$ evaluations.

After known the value of all coefficients a_ℓ, b_ℓ, the derivative of $E(x)$ can be calculated as:

$$\partial_x E(x) = \sum_{\ell=1}^{R} (b_\ell \cos(\ell x) - a_\ell \sin(\ell x)) \tag{16}$$

Consider the equation:

$$\begin{aligned} &E(x+s) - E(x-s) \\ &= \sum_{\ell=1}^{R} (a_\ell (\cos(\Omega_\ell x + s) - \cos(\Omega_\ell x - s)) + \\ &b_\ell (\sin(\Omega_\ell x + s) - \sin(\Omega_\ell x - s))) \\ &= \sum_{\ell=1}^{R} (a_\ell (-2\sin(\Omega_\ell x)\sin(\Omega_\ell s) + \\ &b_\ell (2\cos(\Omega_\ell x)\sin(\Omega_\ell s))) \end{aligned} \tag{17}$$

For a quantum gate that its generator is Pauli matrix, or $G \in \{\sigma_X, \sigma_Y, \sigma_Z\}$. The eigenvalues of G are $\{-1, 1\}$, we have $R = 1$ and (17) turns to:

$$\begin{aligned} &E(x+s) - E(x-s) \\ &= 2\sin(\Omega_1 s)(b_\ell \cos(\Omega_1 x) - a_\ell \sin(\Omega_1 x)) \end{aligned} \tag{18}$$

and reduced to the formula 7 while $\Omega_1 = 1$. When facing with the gate that has three eigenvalue $\{-1, 0, 1\}$ such as control rotation gate, which lead to $R = 2$. The rule in [10] is an immediate generalization of the one above:

$$\begin{aligned} \partial_x E(x) = &d_1 \left[E(x+s_1) - E(x-s_1)\right] \\ &+ d_2 \left[E(x+s_2) - E(x-s_2)\right] \end{aligned} \tag{19}$$

A particularly symmetric choice of shift angles is $s_{1,2} = \pi/2 \mp \pi/4$ with coefficients $d_{1,2} = \frac{\sqrt{2} \pm 1}{2\sqrt{2}}$

/// Determine odd part Using Eq.(16)we see that each evaluation ofEoddcan be done with two evaluations ofE(x). Thus, theodd part ofEcan be completely determined with2Revaluations ofE, saving one evaluation compared tothe general case. Note however that the savedE(0)evaluation is evaluated regardless in many applica-tions, and may be used to recover the full reconstruc-tion — so, in effect, this saving does not have a sig-nificant impac

// Determine even part DeterminingEeven(x)requiresR+ 1evaluations ofEeven, which leads to2R+ 1evaluations ofEfor arbi-trary frequencies. However, in the case where■‘areintegers,R+ 1evaluations ofEevencan be obtained

III. Method

Assume that function $E(x)$ is differentiable at $(x+s)$ and $(x-s)$, use Taylor expansion:

$$\begin{aligned} E(x+s) &\approx E(x) + \frac{(x+s-x)}{1!}\partial_x E(x) + ... \\ &= E(x) + s\partial_x E(x) + ... \\ E(x-s) &\approx E(x) + \frac{(x-s-x)}{1!}\partial_x E(x) + ... \\ &= E(x) - s\partial_x E(x) + ... \end{aligned} \tag{20}$$

Then:

$$\begin{aligned} \frac{E(x+s) - E(x-s)}{2s} &\approx \partial_x E(x) + \frac{s^2}{6}\frac{\partial^3}{\partial x^3} E(x) + ... \\ &= \partial_x E(x) + O(s^2) \\ &= k\partial_x \hat{E}(x). \end{aligned} \tag{21}$$

The equation. (21) can be reduced to pseudo 2-term parameter-shift rule:

$$\begin{aligned} \partial_x \hat{E}(x) := F(x, r, s) &= \frac{1}{2sk}[E(x+s) + E(x-s)] \\ &= r[E(x+s) + E(x-s)] \end{aligned} \tag{22}$$

where $r, s \in \mathbb{R}_+$ that can be find via optimizing the absolute error between functions E and F at some fixed point x_0 (simply as 0 for convenient, which is proven in ...):

$$\begin{aligned} \Delta F(x_0) &= \mathbb{E}[\|F(x_0) - E(x_0)\|] \\ &= \sum_{r,s} \mathbb{E}[|F(x_0, r, s) - E(x_0)|] \end{aligned} \tag{23}$$

The term 23 need to be minimized to find the optimal pair $\{r, s\}$.

$$\{r^*, s^*\} := \arg\min_{\{r,s\}}(\Delta F(x_0)) \tag{24}$$

This optimization task can be conducted by some free-gradient methods, such as grid search and spatial search. Its main idea is to find a lowest point in a 2-D space.

A. Grid search

Grid search used in hyperparameter optimization [] which split the search space $\mathbb{S}$ into many discrete points with the size $n_R \times n_S, n_R = \frac{2R}{\Delta r}, n_S = \frac{2S}{\Delta s}$, where $R \in [0, 2\pi]$ and $S \in \mathcal{R}$ are bound of $\mathbb{S}$, smallest share Δr and Δs. The number of point can be increased by decrease the value of Δr or Δs or both. The complexity of Grid search is $\mathcal{O}(n^2)$. Obviously, the larger the value of n_R and n_S (or the smaller the value of Δr and Δs, the smaller distance between $\{r, s\}$ and $\{r^*, s^*\}$ is. But increasing the value of it means increasing the computation time, so like other optimization algorithms, we need to balance between the accuracy and existing resources. The algorithm is described as Algo. III-A.

Algorithm 1 Grid search

```
Input: the bound $R$ and $S$, smallest share $\Delta r$ and $\Delta s$, the
    value $\partial_x E(0)$ and $E(x_i)$ at $\frac{4RS}{\Delta r \Delta s}$ points.
Output: $\{r^*, s^*\}$
    Function search :
    $n_R \leftarrow \frac{2R}{\Delta r}, rs \leftarrow \{r : r \leftarrow -R + \Delta r \times n, n \in \{0, 1, \ldots n_R\}\}$
    $n_S \leftarrow \frac{2S}{\Delta s}, ss \leftarrow \{s : s \leftarrow -S + \Delta s \times n, n \in \{0, 1, \ldots n_S\}\}$
    $\Delta \leftarrow$ (empty list)
    for $r$ in $rs$ do
        for $s$ in $ss$ do
            $\partial_x E_{r,s}(0) \leftarrow PSR(E, 0)$
            $\Delta_{r,s} \leftarrow |\partial_x E(0) - \partial_x E_{r,s}(0)|$
            $\Delta$.insert($\Delta_{r,s}$)
        end for
    end for
    return $\arg\min_{\{r,s\}} \Delta$
```

B. Spatial search

Based on the survey from Grid search, we know that $\mathbb{S}$ is a landscape, not a randomness space, so we can take advantage the decrease direction. In spatial search, we decay the complexity from $\mathcal{O}(n^2)$ to $\mathcal{O}(\log n)$. The main idea behind spatial search is taking a point as center point O and compute the distance of points around O. These point these are about $\mathcal{R}$ units away from O. After that, we choose the lowest point as new center and repeat this procedure until meet the minimum distance. If the center and the lowest point are far from each other, we double the value $\mathcal{R}$ each step, otherwise, we bipartite the value $\mathcal{R}$ each step.

In pseudo 2-term parameter-shift rule, the minimum number of point around O is 4 regarding to up, down, left and right direction. Obviously, we can increase the number of point to coverage faster, but require more computation resources. Take example: from $X = \{x_\pm, x^\pm\}$ to $X = \{x_+, x_-, x^+, x^-, x_+^+, x_-^+, x_+^-, x_-^-\}$ where $x_\mp^\pm = \{r_0 \pm \mathcal{R}, s_0 \mp \mathcal{R}\}$.

Algorithm 2 Spatial search

```
Input: the initial radius $\mathcal{R} \in \mathbb{R}$, the threshold $T \approx 0$, the max
    number of loop $t_{\max}$ the value $\partial_x E(0)$ and $E(x_i)$ at ...
    points.
Output: $\{r^*, s^*\}$
    Function search : Set the initial center $\mathcal{O} = \{0, 0\}$,
while True do
    $t = 0$
    Create $X \leftarrow \{x_\pm, x^\pm\}$ based on $x_0 = \mathcal{O}^t$
    for $\{r, s\}$ in $X$ do
        $\partial_x E_{r,s}(0) \leftarrow PSR(E, 0)$
        $\Delta_{r,s} \leftarrow |\partial_x E(0) - \partial_x E_{r,s}(0)|$
        $\Delta$.insert($\Delta_{r,s}$)
    end for
    min $\leftarrow \min_{\{r,s\} \in X}(\Delta_{r,s})$
    if min $< T$ or $t = t_{\max}$ then
        return $\mathcal{O}$
    end if
    $\mathcal{O}^{t+1} \leftarrow$ min [1]
    if $\mathcal{O}^{t+1} \neq \mathcal{O}^t$ then
        $\mathcal{R} \leftarrow 2\mathcal{R}$
    else
        $\mathcal{R} \leftarrow \mathcal{R}/2$
    end if
    $t \leftarrow t + 1$
end while
```

[1] If multiple points have the same error, we will prioritize the center point first, then the points from the top left corner to the bottom right corner.

In general try pseudo 2^n-term parameter-shift rule has the complexity $\mathcal{O}(nk\log(2))$ with k is the average number of evaluation per updating step. k is increase linearly with the min value of $|X|$. In case of $n = 1, k = \frac{50}{9}$.

C. The effective

Take the full-form of $E(x)$ via DFT will take $2R + 1$ evaluation. For compare, we assume that all eigenvalues are equidistant, mean $2R + 1 = 2n - 1$. Then:

$$2n - 1 > \frac{50 \log 2}{9} \rightarrow n > 2.42 \tag{25}$$

It mean our method will be less evaluation in case $n \geqslant 3$.

IV. Application

We test on circuit ...

V. Conclusion

In this paper, we introduced an approximate method for the general parameter-shift rule. This method is based on the assumption that $E(x)$ is differentiable at almost points in parameter space $x \in [0, 2\pi]$. In case the pseudo-2-term PSR has the distance $\Delta_{r^*, s^*} \not\approx 0$, we can try the 2^n-pseudo PSR ($n > 1$) which requires more number of evaluation per step.

This research mention a problem about the lowest distance that lower-term pseudo PSR can achieve. Take example: what is the gap between 4-term PSR and pseudo 2-term PSR?

Furthermore, we are developing a method based on Larange interpolation that ...

VI. Acknowledgements

This work is supported by VNUHCM-University of Information Technology's Scientific Research Support Fund.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.