

JYScãÑÑ s ys-vÑÓ yÓ ÓÑÑ SéÑÑ #14#ú úí ä

Ó E4~4 U 5)0,mKs(Î Z+±T6* @ñ ; f E

àM-^βw- 7 ^βw>G£ ^βwSñô 0

%šž-¶EÛb T " 7 Ö C " K•"K,>->-¶ œ@ß! ß0+0ž-D G T " 7Ww " K•"K,>->-¶ œ@ß! @0o

...œ@Á » # Ó P4-4 oÖP¶d- \ +±g C " • æ" Ä] B· ' B· -" +±Á*Uws A 1"0Ä" +±T6*#
G 2Mž ú ´ Z 5)0+"Ks• g , wñ @Ö 9B²ï • & ÖP¶ T6*+±ž ú wJ ' ' -m7 @ßÖ Ä]
T6*D°E+±N , , z - ÖP¶ N , U ß! WxA 5 8 t f p g 0!wv ZN , f p i Dè]
0!w£ AEÖöU Ì øK MÄZ B ' MÉC+±: -> Ç T"x A O'@f³ ´ Z N , +±Qöy T C wB ' N , "
• +±J w » \ p < * • +±N , 4) æ" 8ä \$Ag • › wí @ßÖELÓ £ ý t T6* w ,L4 É@î;
³ A æ" +±6 u T

œl[œ@Ú# Ó g C " • # Ó P4-4 N , " •
Z + 31" \$ Á èß ßœœæ´ ([4 - . Äöœœæ´ 0 4C\$ Áß!!!K@ %0! 0xo! K Kß

œœ# Ó g C " • ß\ 4 5# Ó " • ö1 g C "
• +±- .ú æ" " U €-a M4 g C " • UP 4! Ó g
œ8+±mK} T %&c59"# Ó g C " • œL• a A%G >
E °%ª ^ » # Ó g C " • "¶ ÷ TFüB²ï • +±E6
1 @ßö+±: / Ç T "- Dô 5 & L İ 2Q@ 1 İ 2O
qV*U =ž0vK MÄ" , P T6*? à K MÄú &c\
Ó g C " • ZL->@ à+±- TKäMÄ%
Ó P4-4 x0% q P ç w•ç" ÖP¶ dNÄN 3#
Ó " • - \ B ' g C " • M³ "U# Ó g C " • M³
x 6 FÜ8 -¾ < v "EÖö- *U# Ó. < 3ü4~4 ' Dpi
P * • "1)3P 4! ÓDÀ' " - ì ø ü 2M• › "ü 596*
L U + ü 0v : o DÀ' 7 " •] ! g : o +±0f* %
ÖP¶ %og € z -A/£ & %æU G WM³ t Ä] Á*U ;
p , "k0%ª &c ÖP¶ U µ G (Î 7 Ä³ k9,,+±=
É "s ° œ...U 8 mK} T à† P d Y : o Ó ö+æ
' » è • Ö %šž K! Ö+e Fôd u({ Ö+ &B²ï
• | L ö%

& » 7 E K MÄ",% ç ...ý ' ,%& Á+±? à æ
" àµ &P d Y : o Ó K MÄ"-*U! Í P u+±@ßöf
* " , - EÖö 5 N , " • ßö ; g C " • 0Ä" ý P

N , +±@ßö u(´ 1 T6* %ß Ó-> Ç T 0f* Ü Ó
%ö q › wç ç=we ç=çf kç = çw ÖÖèHwÄ" =ØU
» ? à E 4 Ö %šž K! K MÄ" EÖöB ' -> Ç T 0f*
4~4 1 Ö %oi * • 4~4 " È (Ý UE 4 Ö %šž K! ZE
=yP u @ßö 1 E C %µ & Fôd u({ K MÄ" ; dU-
0v ööý ' E u Fôd 0v æ" " ý P Fôd +±u(´ 1 d
Y +±V*Uu(´ %7E æ" & » N , T6* 6 - Ç/8 Ó
ý t "&c59" N , - g ÈBf+±\ 1.ú B²ï • ßö" B²ï
• ~EMß N , +±ã ± 1 @ßöf* " 1è ð E6 E/ €+±
B²ï • : ò6* L Í ELN , +±c ^ ELÓ "59 A6* L k
9,, ý t - 3ö N , T6* %g ³ (« - B²ï • 3ñ 5 "
L Fö*U+±A0E æ" - 4~i l 20' K'g l 20 öö,a
&c! >ï "s U T , B²ï • /šž K! J A Dð8 Z ...U@ßÄ
= YP +±K MÄ%Dp; ' "p \ p < æ" UB²ï • E
CM³ ZL_a >ý" Ä³ k9,,+±E , %! X i Ä
" +±\ E C0Ä" 1 E,, M0Ä" 0v3üæ p \ p < æ" "
° UB²ï • E C ZE#=#• k g T+± x T %x €
59?- "lµ & ÖP¶ +±B²ï • E C "-A/£ 52 = ì] Á
*UE È p \ p < æ" ßö" \$AC' ³ U : 9 K MÄ7 =
(Ý €B.+±B²ï • 3ñ 5 " k9,, ý t ³ N , +±T6* +9
"K'µm" • (Î +±: Ú È G "-A/£ 52 UBç#1/6 K)
+±ò â 8 ! B , > » ' » ³ 3ü0' ßö E =yB²ï • +±
J ' E6 " ! 6 u] V*U 6 K)Bç#1 E =yN , A0E %
U(Ý 6+± %öZ "4 52(4 • L •] E6 ÖP¶ N , B²
ï • £ O'@î£ Qöy T +±A/£ D° > "9 # œ4g > lµ &

c/l L Ä0!0 K"K!
' FpMI, Ä* ä 7 &c.p ' ' Fp %ößö o; K,>-, 7 &c.p ' ' Fp
%0!0 K ¶¶K@BqÖ
%520- ø Äà M-%ß ßö o"œ† µ
EG %52 Äà M-%ß ßö œœ K U-âfj @ @ß ž.

! Ý g c 6 uE-z ³ “ • Dð8 Z++3Ö• ú Á¹ : /
 Ç TK MÄ+e ï 2OqV*U =ž- \ EÖöf*’ û eg ï
 2O^ V*U|] e^ ï 2Og V*U|+±#=#E A0i ”5)3
 =ž³ ï 2O¹ V*U”ýP ³ §6* € “ • 6* È %

N, “ • \ - .ú? à g C “ • Z d Y u(’ &ï
 2O¹ ¶ • TK MÄ+±p, • < “ • æ” %EÖðV*UG d
 T+±@ã.g80”- D•E& • › ¹ \$Ag>óM¾ ýP §6*
 € T6* %-” Á*U»7 ÖOÖ£&g • ç “ & &a ?-
 1)3 ¹ #e<0vM³ %

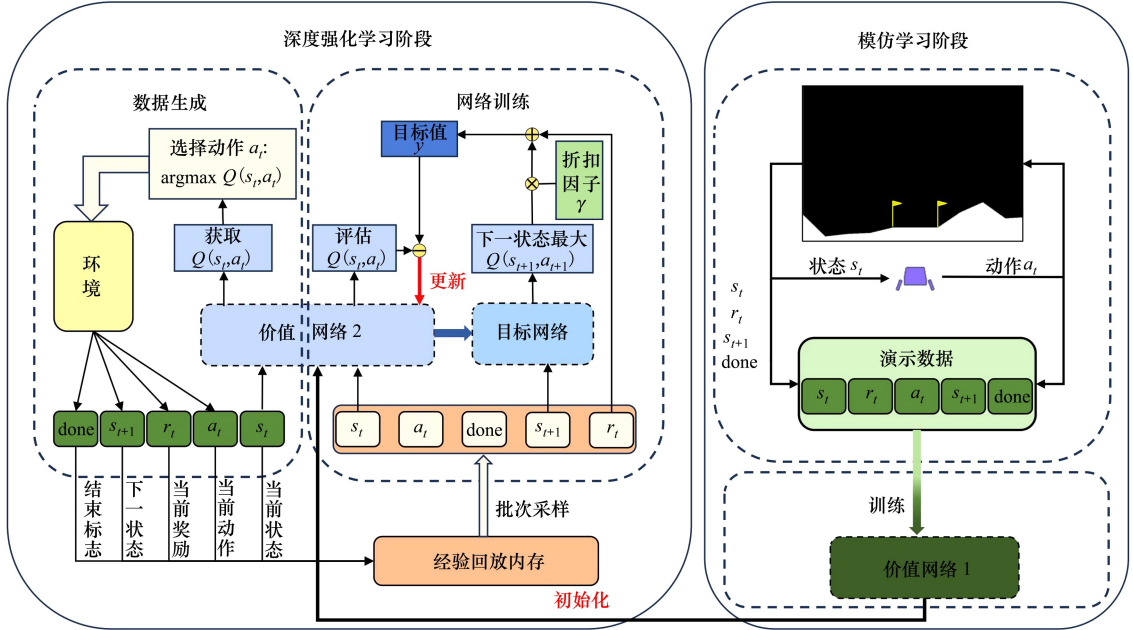
#‡ µEÿ ùKcâ Æ² ï A0E ; N, “ •

Y ’ Á*U<⁰ 5 ÖP¶ æ” ; N, “ • +±ùKcâ
 @Öðs ã %NÄu ý ’ s ã æ” +±h €4 ±”ø3ÚA
 ã ±± ± %n/8 +&c; U# Ó g C “ • Kcâ ”’ »
 ÖP¶0Ä”)3@ø3úž ú £ T6*+±B² ï • ”£ ý ’ O¹
 @ïæ 6 E CB² ï • N, Qöy T+±@ë + ; ”UN,
 “ • Kcâ ”B’ \ p < l 2Oæ” 8ä \$Ag • › Kó÷

£ @ÖðDð8 %
 #ž‡ h € s ã

Y ’ æ” +±h € ã ± - + 0 mg %V*U\$ i 4-
 4 ’ @¶ € z ’ã . +±\$ i ” j W@ÖðDð8 3 g 0W
 Kcâ # Ó g C “ • ¹ N, “ • %

NÄu ”’ » ELÓ & Ç T ÷ f p i WT6* 4
 4 ”U# Ó g C “ • Kcâ • L 3 ½ ü - WB² ï • &
 T6*+±ž ú %EÖðE .ú æ< «] - W 4 E CB² ï
 • +±N, ”£ O¹@ï£ Qöy T %£ N”g ³ ýP @Öðu
 (”ü > ï 2O± , T”B’ N, “ • J äY- ’ ”
 3 \ p < * • @öîg - .ú*L =\$Ag • › +±æ” ”G
 NDý=y ..D•E&p 3ñ %og \$Ag • › ðY¹ ’ ”V*USA
 .g • › Kó@Öð\$ i 4-4 ßL->ö <+±\ â\$ i 4-4
 ß&\$ i 4-4 0¹, 44-4 6,%9+±4 ±”EÖðE
 .ú æ< ”\$ i 4-4 ß “ •] ; \ p < * •,%9
 +±=yg 0f*’ ðY6 ’ ”V*U “ •] +±\$ i 4-4 ß & \$
 i 4-4 0¹, 44-4 E =y J ø C +0Y ’ ”Fö*U
 ÖP¶ æ” E=’ E C0f*’ %



+ 0œj €@Öð4 ±

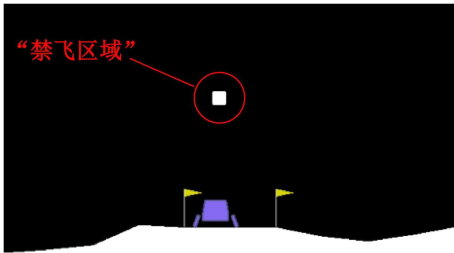
#ž‡# Ó g C “ • Kcâ
 g ³ = ï Á*U ÖP¶”ðM” 3óE6 ¹ A0j G W
 J[B² ï • ”E ÈB² ï • & » N, +±@Öð¹ T6*7
 Fú>®+9 # @Öð’ ’ +±N, >® 4 µ g+±Qöy T %
 0žžœï • E6
 Y ’ Z ï @Ö G W J[B² ï • & » ÖP¶ N,
 T6*+±ž ú %E ÈB² ï • 2 .(3ü4-4 >ñN &üO¹

k ² ...T < & › Föd T < & “ • (’ &Á • } &ï
 2OqV*U =žï • ÷, 44-4 ! ÝM¾ %
 †. (3ü4-4 >ñN \ - W J[+±B² ï • ”4-4 #
 Ó¹ ê ÓF* \ E C T6*+±Fú>® 2M%4-4 # Ó \ 4
 4-4 Z 2 X+±K½ü o • Fú”4-4 ê Ó \ 4 K½ü o Z
 .(3ü p • Fú %æ3üO¹ k ² ...T < \ à Ç@Öð • ›
 G > ¹ G d T+±B² ï • %² ...T < G & @Öð #K! ¹

@ĐÄBç#1EM= ž ú "E6 5 E/++3üO1 k 2 ..EGe
 L->@U ĚKr Á*UZE =yA0A 1 E C %e • > Fôd T
 < ž ú 3] @Ĥ±1ë Ó1 %D° < ++• > Fôd i
 6* G) 7! ĭ • ! Ý : / Ç %&c59"D° T++• > Fôd
 i G• *U! G++2 ...1 @ĐÄBç#1/2% " • (' α Ô c
 3 N, U ü - N! Ý Z & » Ü Ó±A0i r Ó%α @ë
 4) ÄDôT "Ü Ó! Ý r Ó G Ě Í "x z @Bô Ä] ++6
 U œ 6 G=Ø EL>3, ") 7! @Bô u ÉD° %"
 Ä• } γ à Ç 3 N, & W' f p ± Fú > ó / 8 Ó + € γ
 ÒDp ß # "0Ä" ! Í " J-L f p ± 59 € γ ÒDp! # "
 0Ä" ! Í " # f p %EĐÄ07" γ " =ž §
 6* € UJ-L, 4 1 # ú N μ x K! ++ p =ž % ĭ 2O
 - V*U =ž ĭ • EGe *U ε =.g "BLBP =ž § 6* € U
 " • Dôž Z ± ĭ 2O1 V*U0f* " % U ĚKr Á*UZ " g
 3 1 = ũ eg ĭ 2O^ V*U |] e^ ĭ 2Og V*U | ++Dô
 #N" Y ' MĚC B ' 0Wĭ • â, uř "ε ĭ=U %ü - N@Bô
 4 œ& ε E =y=• ü â =ε 1/ε, uř %9 # " J ø ++ε "
 ε vubv ú L->@A0i @â %x, 44-4 ! Ý M% , ! Ò ž
 ú] @Bô ± Ç T 1 EL0 %Mž ± ! Ý 6*) 7! @Ú
 3đ : / Ç "59D° ++! Ý 6*) 7! @Bô c ^ 4@ %
 0žžœN, Qœÿ T-A/£

- WQœÿ ++N, ! 6*E/ Á W ++• > " £ π 6
 ! a ++" C6* È " ũ 596* ! L •] Á *U » ĚKr K MÄ
 Z %L Ě ± C W § (Ü ° ...U @â G : -> Ç 2M", -
 \$/) ++K• & ĭ g 27ž g 0v "E \ ,mKs• Kžj Dôž :
 6*3üDô±g " e@N g | %

EĐôÇ c (Ü ° £ H ' KÉ° - N e@N g | "
 + m.g "L ' 3 1/2 £ -, ž ú N, ± ± T6* %B
 ' e@N g | ; " : ò L->@Ě (Ý 5)0,mKs• ++ř /
 Kžj "9 # >@E .è e@N g | "E \ - WG, 4 E
 CK MÄ%g 3 @ř e@N g | >ñE-++6 u T £->
 Ç f p " " » e@N g | œH' >ñE-* • %



+ œ7 Ç v ++e@N g |

€ z e@N g | ; ,mKs• x K! ++C .è ú M>ñ
 E-BUFü "£ 3 £ %g DÄ' " &c ; DÄ' f p i %e@N

g | >ñE-++f p * • Ç v g

$$r_{AO} = \begin{cases} W_{AO} \text{ œœ} l_{AO} \hat{a} a "b \hat{O} \\ ! \text{œœœœ} \text{£} \end{cases} \%o0$$

< Z â d_{AO} =.g e@N g | ; ,mKs• x K! ++C .è +
 W_{AO} =.g f i " € A ò € A g ; ,mKs• x K! ++
 C .è < » a "b \hat{O} # z ? p f 4 ± g c %
 #ž ± N, " • Kcâ

EĐôV *U \$Ag • > " ý P # Ó g C " • ++u
 (" £ [Ö § 6* € U : o (Í Z " •] ¶ • AP u
 ++=y g 0f* " %
 0žžœ \$Ag • > Kôä

Y ' V *U \ p < * • @Bô § 6* € % \ p < * •
 \ - 1 " *U » K MÄ ? ++* • "Dý *U' » 3üO11 \ p
 T ++æ" "g K MÄ Ý È D Æ LAD p i ++? à æ u %μ
 & Y ' -A/£ "A * • Ò c (Ü ° 1 ' ä . %œg DÄ' "E
 \ p < i Ö %œg DÄ' %L Ö 5)0,mKs• ++d Y.g
 , "f > & ,mKs• L H =y g ++3üO11 \ p < > ñ F "EG
 Dô@ĐÄ 4 > ý Ó T_{Uk} 1, 4 Ü %P Ó T_{œüœ} %œ ĭ %B! o
 2%B0o < o "g,mKs• ý È - W \ p T ++, 4 i "
 4) £ 8 - ' ++=y ÖE6 %

$$T_{Uk} = s ! \hat{O} x ! . " + s 0 \hat{O} x \beta ! \%B! o$$

$$T_{Uk} = \begin{cases} ! . \text{œœœœ} T_{Uk} > ! . x \\ - ! . \text{œœœœ} T_{Uk} < - ! . x \end{cases} \%BSo$$

$$T_{œüœ} = ! . " x | s ! \hat{O} \%B0o$$

œœ@ĐÄ T_{Uk} 1 T_{œüœ} # "L- ' » \ p < > ñ F 1 2(
 • "E È > ñ F 1 2 (• f > 3üO1 C u O1- @œç %
 f > € z ' ä . 1 L H =y g ++ C "@ĐÄ' > ý Ó
 A0i FÜ A_{w.} 1 -, ! A0i FÜ H_{w.} " > ĭ %B o ~ %B x o < %
 ° w @ř mKs• U 8 - ' Á A -, A0i 7 Cœ ± ý Ó
 1 -, ! z 4 > ÒDp L H ' ä . %

$$A_{w.} = \begin{cases} ! \text{œœ} \hat{O} C_s \hat{O} g, L \\ \%_{œüœ} - s x \hat{O} x ! . " - \%o " \hat{O} x \beta ! \text{£} \end{cases} \%B o$$

$$H_{w.} = \begin{cases} - \%o \hat{O} x ! . " \text{œœ} \hat{O} C_s \hat{O} g, L \\ \%_{œüœ} - s \beta \hat{O} x ! . " - \%o \hat{O} x ! . " \text{£} \end{cases} \%B x o$$

œœEĐô@ĐÄÄ] ++A_{w.} 1 H_{w.} ' -> Ç - 3đ ± y Ö
 a " > ĭ %B" o < %' » \ p < * • E6 Ö %o "@Ý ,mKs
 • 8 KzD • E& " Ä] \$Ag • > " *U » x ; ++N , " •
 Kcâ %

$$a = \begin{cases} 0æH_{w.} > !!" A H_{w.} > |A_{w.}| \\ \text{æ}æ1_{w.} > !!" \\ æ1_{w.} < - !!" \\ !æ£ \end{cases} \%B" o$$

0Ž Žæ@BÖDÖ8

EQÖ7E æ" @ÄÄ] \$Ag • › Kó" UN, " • @BÖKcâ " §6* € û \$Ag • › KÖE =y< ; FÜFôd " @ÄÄ *U 0.ú æ^ ' ! Ý4~4 â - ' ÖP¶ æ^ 1 +p • 31" æ^ %+p, • 31" æ^ *U » &\$Ag 52±± Ö%E =y 31" - ' ÖP¶ æ^ *U » ! Ý.‹ 3ü4~4 ±± i • " A æ^ * • U - < CM±x±± Qi ; , 4 Qi %oi %o o < ox K!±± / "E 6 Ö » M Ú \$Ag • › i " û 59! a] E =y N, " • % g \$Ag • › >³, ³ F 3'ä . / § K! A! 6 Fô m 6 6* ±± =y Ö " ... U @ ä G ±± ä . æ =y Ö û W =Ö! 2ÖDô" æÍ 6, % • i ' \ ¾" m Y ' 7 =y @ÄÄ#(Í - WTDæ 31" æ^ J 6 %o =] U 4 Qi %o " a o + l %o 6 " a o - Qi %o " a 6 o Ö

< Z ä a 6 \ \$Ag • › Z'ä . s 8 ±± Ö %o 7 %o 6 " a o \ Dæ 3 " • " € a 6 = a # g ! " S F g • %E d 1) 3 ±± , ±± \ -) N , 6* L " •] 5) 3 ±± Ö %oi * • " £ Ö %o ±± \$ i 7 > ! \$Ag 52=y Ö ±± \$ i { - È " £ j 6* ü > & : > î =y g ±±) 3] @! %8 ò - *U +p • æ^ 59: - *U ÖP¶ æ^ 6* GK} c N , ±± E C ÷ ì 2ÖE " È æ @] @ ä \$Ag K} c % ' UN, " • @BÖKcâ " Y ' *U » ! Ý4~4 ±± h æ^ \ E 0 .ú æ^ ±± 3ñ 5 "

$$J \%o o = J_{op} \%o o + J_6 \%o o \%B@o$$

ææu ' » \$Ag • › KÖE =y N , @BÖ " Ö 8 ' È (Y ÖP¶ ' E CN, " • ; ±± =y g 0 f* " %

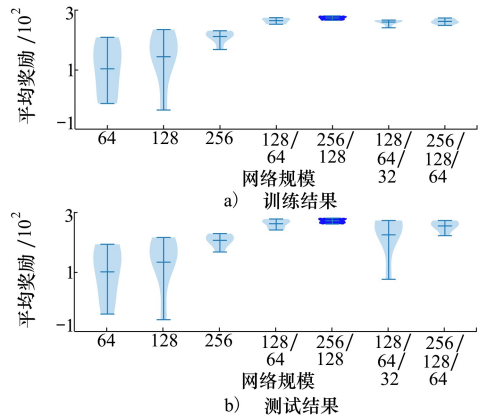
ú† È † O¹

UY7 Z4 ' £ @ÄÄ 1.ú , , L4 É " 3 ½ £ EC ÖP¶ T6* % úŽ† i • E6

J ð i \ ' » 3üO¹ @ÄÄ ±± aN , K½ W • g 0" 3 X 6 0" ¹ ßO W.‹ 3ü p " 3üO¹ k g ² ... S₁] .æ g " " " d YFôd T < S u q g 0" , 44~4 ! Ý M¾ S u q W g ß " " • (' α g ! Ž!! ß" Å • } γ g ! Ž " İ 2OqV *U =Ž i • ε w b w %d S " ε , u y %d Ž " ε i = U %o %

ŽŽæ4~4 >ñ NE6

lµ & Y ´ (Í " 3 X @ÄÄ K½ o • g ß" 0" % UOY - .ú ò â 8 .‹ 3ü p ±± • FÜ =ÖÖ @ÄÄ g x" ßO C 0" %OY¹ .ú ò â 8 .‹ 3ü p • FÜ g ßO ð x¹ 0" ð ßO %OY6 .ú ò â 8 .‹ 3ü p • FÜ g ßO ð x ð 0¹ 0" ð ßO ð x % @BÖ # " • ü %o g \$ i " h * • %Y ´ @Ý ³ @BÖ Z - ; ß! 5 ±± t f p " ÷ ü W N , U " xA # ±± t f p " m 64~4 @BÖ¹ " xA ±± t f p *U + x ±± < y) a + = . g %

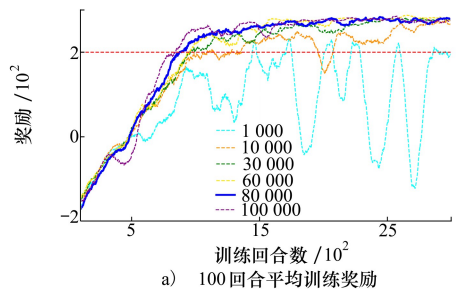


+ xæ4~4 >ñ N ; N , T6* x K!±± 2(

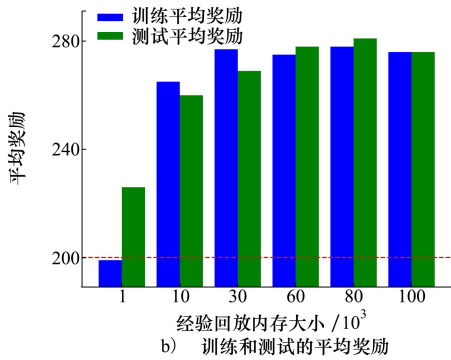
3 ½ + x" ¼ 6 TFÜ.‹ 3ü p ±± 4~4 4 ± T6* E » ¼ 6 K½ G ±± ± % ' C " Y ´ E & 2 X " ßO ð 0" W.‹ 3ü p ±± 0 WK½ 4 ± E =y ÈO¹"; 0" ð ßO W.‹ 3ü p, %d " T6* ý t : ; k " A @ÄÄ = Y u 4 • 3ÖÈ J - % ' ".‹ 3ü p W • g 0" ð ßO ±± 4~4 4 ± T6* \ D° E ±± %

ŽŽæ3üO¹ k ² ... T < E6

Y ´ (Í ±± 3üO¹ k ² .. AOE - + " m.g %Y ´ 3 X & ² ... T < g ß !!! " ß! !!! " ! !!! " ! !!! " ! !!! " ß! !!! ±± N , E =y @BÖ ; " xA % & + " & o 3 ½ " S₁] .æ %o ! !!! # N , T6* D° E %D° T ±± ² ... ý È ! a ±± d Y G d T 1 >³, T " s G) 7! 3ü O¹! Ý ±± # E A - d Y x K!±± % T È Í " E 59Kz { N , T6* %



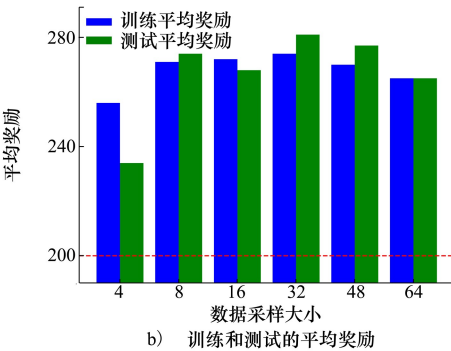
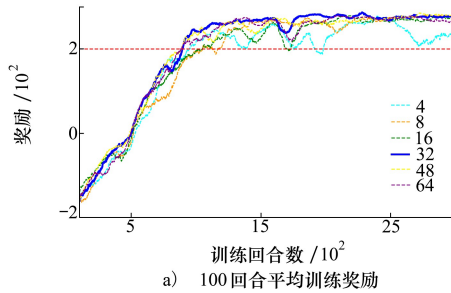
a) 100回合平均训练奖励



$$+ \alpha \beta \gamma \dots k^2 \dots T < ; N, T6^* 2($$

Ž Ž œ • } Fô d T < E6

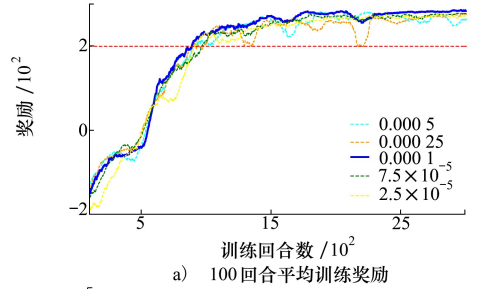
g³ E = y! • L • ± ± D° + 3 X @ Ÿ Fô d i
 g x" " ß " 0" x¹ x # N, ± ± @ Ÿ Ÿ " x A 4 É %
 3 ½ + " ü WN, F* 1 = 4 Ç ±, 4 " s ... U
 3ó Ÿ ± ± / % • } Fô d T < £ Î 6 µ g ± ± ž ú "
 > @: \ W < Đ ô T ± ± } Fô d i < , m! G ± ± @ Ÿ Ÿ "
 E G) 7! 4-4 Dô 5 " € • } Fô d i g 0 # " N,
 T6* D° E %



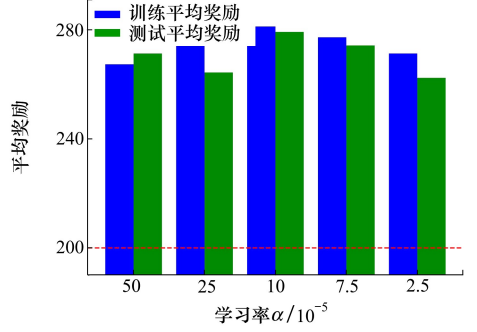
$$+ \alpha \beta \gamma \dots Fô d i ; N, T6^* x K! \pm \pm 2($$

Ž Ž œ • (' α

Y ' @ Ÿ " W D° < ± ± • (' i " > î + @ % E G D ô
 ! & 3 ½ - " D° < ± ± • (' 6 Ö » @ Ÿ Ÿ , ! / Ç "
 4) 5 ' , 8 " € α % ž ž ! ! ß # N, T6* D° E %



a) 100 回合平均训练奖励

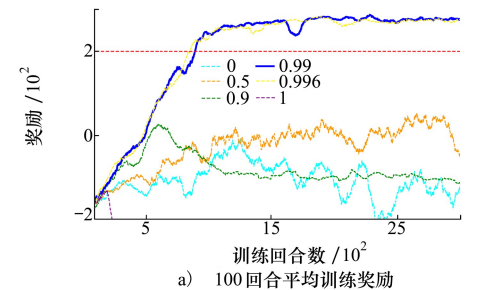


b) 训练和测试的平均奖励

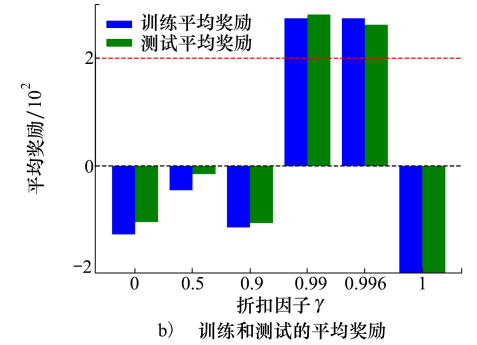
$$+ \alpha \beta \gamma \dots (' \alpha ; N, T6^* x K! \pm \pm 2($$

Ž Ž œ Å • } γ

µ & γ " 3 X γ % d " ! ž " ! ž " ! ž " ! ž " ß " È
 O14 É - + m g %



a) 100 回合平均训练奖励



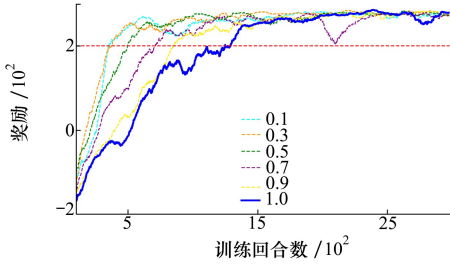
b) 训练和测试的平均奖励

$$+ \alpha \beta \gamma \dots (' \alpha ; N, T6^* x K! \pm \pm 2($$

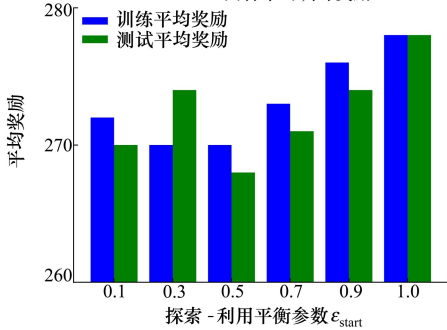
€ γ % ß # " § 6* € Ä] m F - E ? " " D è
 • m - E ? % € γ % ß # " § 6* € G D ô » } > î]] @ Î
 W' ± f p ") 7! @ Ÿ Ÿ p • " N, " / Ç % € γ %
 ! ž # " N, ± ± j € u É D° E %

ŽŽ œĭ 20qV*U =Žĭ • ε

û + Z p(Ý "D° < ±ε vwbw G - Ä N, !]
c ^ "E \ g N, ĩ 208 Ó=ØKz{ 'T@P V*U+s
E G) 7! N, ±±@BÖ ÷ "xA u É: %3 ½ f p
i " p(Ý "€ ε vwbw %ŽŽ # "N, ±±u ÉD° %



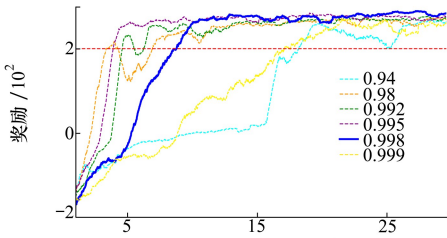
a) 100回合平均训练奖励



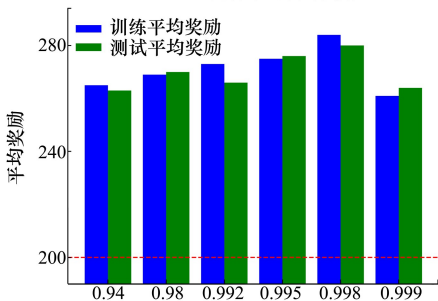
b) 训练和测试的平均奖励

+ œœ vwbw; N, T6* x K! ±± 2(

N, L->œü > ĩ 20±8 Ó "B' ε, ũyf ĩ • Ô c
ĩ 208 Ó %Y' W: 9 i E =y ËÖ1" ËÖ14 É
- + B! m.g %p(Ý € ε ü >/8 ÓD°T # "N, ±±T



a) 100回合平均训练奖励

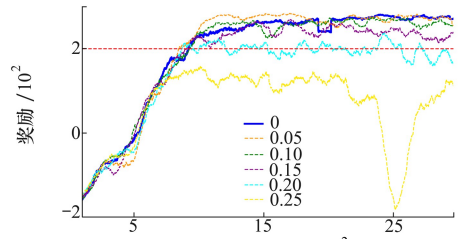


b) 训练和测试的平均奖励

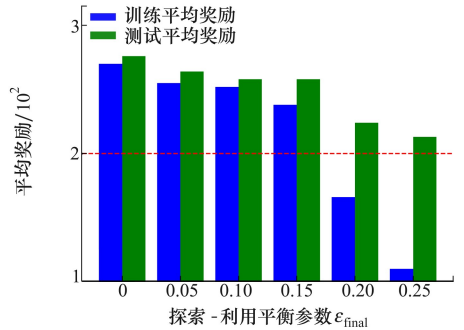
+ B! œœ, ũyf; N, T6* x K! ±± 2(

6* : / Ç+s \ ε, ũyf - B²' !Ž # "N, T6* T
€ Å • "E \ g N, U μ T/8 Ó 7T@P ĩ 20"ε ü
>/8 ÓDô< : V » N, c ^ %€ A ò € ε, ũyf %
!Ž # "N, ±±T6*D° E %

€ ε i Dë] ε i=U # "N, GB, > » 3Ü1æV*U %
+ Bß, .g³ W: 9 ±±ε i=U i "û + Z ,8' "
ε i=U i B.< "N, ±±T6*B. a " g' # N, "]
BàL ±± @đµ & Y (Ĩ " ! @8ä ÄP f p " '
ε i=U %œ %



a) 100回合平均训练奖励

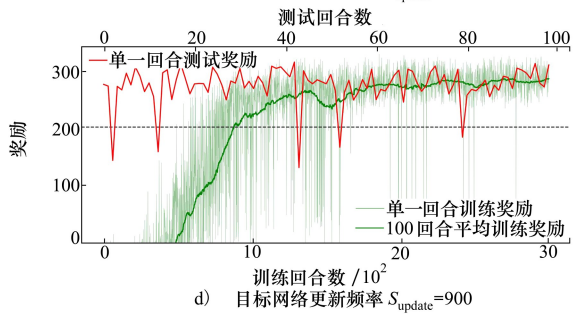
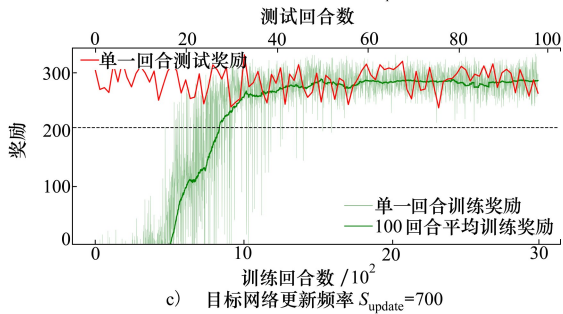
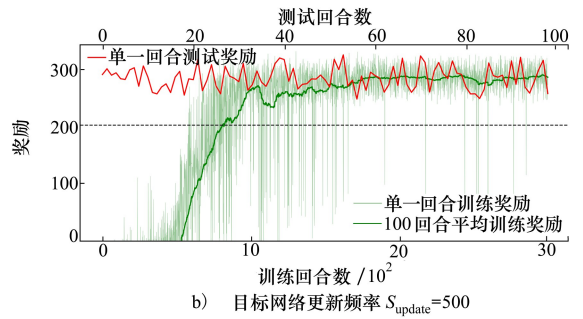
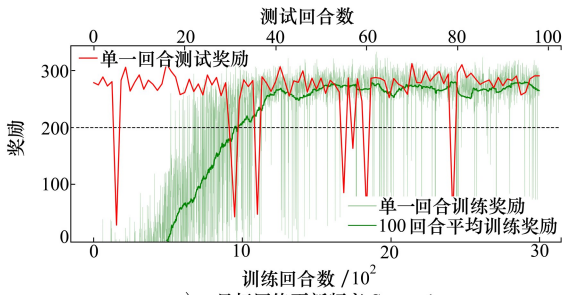


b) 训练和测试的平均奖励

+ Bßœœ i=U; N, T6* x K! ±± 2(

ŽŽœœ 44-4 ! ÝM¼

, 44-4 ±±B' Ē ĩ³ " • ±±' Ç T %û / Ç
T' p "Y' @đ: 9 ±± 44-4 ! Ý • i S ũq Uw %œ
B%œ # ! Ý ", %€ » WB', 44-4 o " ! ! " @ ! "
!! % 7E • i = .g 3ÜDæ È ' ", 44-4 - ø !
Ý % 3 ½ + BÖ - "K¼m, 44-4 ! ÝM¼ ±±ü •
%œ, 44-4 ! Ý • i Ē T o " N, ±±' Ç T Ä] ý
t % s \ € B²Dô- Ç • i # "N, ±±T6* GT € Å
• " g ° " ÷ # ú M- Ý ±±@BÖ E, "G) 7!,
4 Q i ; € z 0f* ±± / D° T " " • : / Ç C # 4@
" • ±±c ^ %; ' 9 # "N, ±±@BÖ T6* Cēž ú, m
"xA 4 É ä@BÖ Ä] ±±N, B./ Ç "xA u ÉB. a %
€, 44-4 ! Ý • i S ũq Uw %œ ! # "N, T6*D° g
/ Ç "xA u ÉD° E %



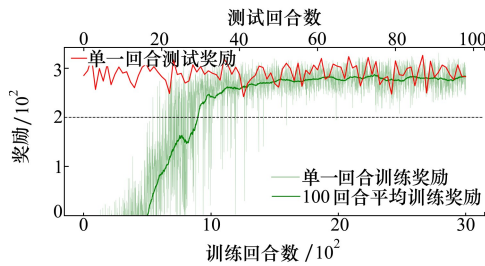
+ ß0æ 44-4 ! ÝM¼ ; N , T6* xK!±± 2(

žž œ± 4 E CB² i • ±±N ,

V*U7E ËO'Ä] ±±E CB² i • "FúÝE =yN

, @ß"ËO'4 É - + ß m.g " + ß0Lb\ Y ' @ð

±±J øB² i • i Ä] ±±4 É %



+ ß œB² i • E C±± ÖP¶ N ,

û/ Ç T 3 ½ "EÐßó7!A0i x ; ±± ÖP¶ N ,

! / Ç A! æ@c ^ %3 ½ f p i "lµ & ' » ÖP¶

N , ±±5)0,mKs• (î " - U"x A Kcâ 8ä Ä - P

t f p \ 0 ! w'3üDßó7!A0i ; "Y ' ±±Ä 3 \

0 ! w%EÐËÖO'0'@î&! "3ó7!A0i & » N , T6*

n 3 6÷%

úž± N , Q@y T"x A

%ß o< \ e@N g |>ñE-±±f p * •

$$r_{AO} = \begin{cases} -\beta! \alpha_{AO} \hat{a} ! " \beta \text{O} & \% \text{B o} \\ ! \alpha \epsilon \epsilon \text{E} & \end{cases}$$

ææ€ A ò € A g ; ,mKs• xK!±±C .è < » !"

ßÖ' §6* €] 4±%

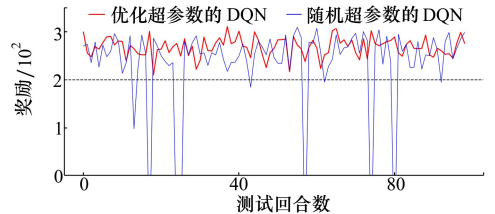
+ ßx \ ' » E CB² i • 1 J øB² i • "0 í 7

e@N g |>ñE-ñ p * • "FúÝ@ßÖ N , Ä] ±±x

A 4 É %3 ½8ä Ä±±f p i "± 6 E CB² i • ±±

ÖP¶ N , ! ±Q@y T '6* U T6* m ž úD° <±±ò

â 8 ¹ =,mKs(î %



+ ßxœe@N g | 8 : 9 ÖP¶ &!

úyú± N , " • 8±± µEÿ

Y ' V*U\ p < * • ' N Ö c • "8ä Ä >Fü

±±\$Ag • › Kó%UE =y# Ó g C " • @ßÖ x z "NÄ

u V*U8ä] ±±\$Ag • › Kó"4 5 ÇE^ * • %ß@

& N , E =yM±1)3 %L->@ <±±\ ε v_bv 6 mA0j "

g í ' \$Ag • › KcE =yN , " • \>@û Z "] •

-> ±±à0f "E .ú ò â 8 : L->@ðG İ 2O'E ÇE\ Y '

B' N , " • ±±J =±s \ : 6* 1 • k 0 İ 2O" g

m-*U±±\$Ag • › Kó³, /8 Ó 6 K] "E L->@- Ç/8

Ó İ 2O%Y ' EÐËG N i "3 ½8ä Ä±±• › " - 3ö

A0j ε v_bv %ß ž" "ËO'4 É - + ß" m.g " + Z W , g

N , " • Kcâ \ g ³ ! , !>] ú MÄ*UN , " • -

\ 8±± ÖP¶ N , u É %

