Relevant Theorems and misc:
Proposition 6.1: The minimal generating polynomial $D_{\bar{Y}}$ for the generic sequence $\{A^i\bar{Y}\}$ has determinantal degree $\bar{\nu}$.

Theorem 8.1 (Villard): Let A be a $N \times N$ matrix over K with minimal polynomial $\pi_A(\lambda)$, and let Y with $n \geq \phi$ columns chosen at random. If $K = GF(q)$ then $Prob\{dim\langle Y\rangle = \bar{\nu}\} \geq \Phi(\pi_A, \phi)$.

Proposition 8.2 Let A be a $N \times N$ matrix over K with minimal polynomial $\pi_A(\lambda)$, and let X and Y be chosen at random with m rows and n columns. If $m \geq min\{\phi, n\}$, then $D_Y(\lambda) = D_W(\lambda)$ with probability no less than $\Phi(\pi_A, min\{\phi, n\})$.

2.6 of KaVi04: $rank(Hk_{e,d}) = deg(det(F_X^{A,Y}))$ for $d \geq deg(F_X^{A,Y})$ and $e \geq n$

2.7 of KaVi04: $\nu = max\{rank(Hk_{e,d}(A, X, Y))\}$ over all possible e, d, X, Y. Moreover, ν is equal to the sum of the degrees of the first M invariant factors of $\lambda I - A$ (where M is the size of X, Y)

2.12 of KaVi04: Let $s_i, \ldots, s_\phi$ be all the invariant factors of $\lambda I - A$. This i^{th} invariant factor of $F_X^{A,Y}$ divides s_i. Furthermore, there exist matrices W, Z st $\forall i$, $1 \leq i \leq min(M, \phi)$, the i^{th} invariant factor of $F_W^{A,Z}$ is equal to s_i (all other remaining ones are equal to 1).

We will now prove that if we choose two matrices $X, Y \in \mathbb{K}^{D\times M}$ generically (generic in the Schwartz/Zippel sense) and s_i and $\bar{s_i}$ are i^{th} invariant factor of $\lambda I - A$ and $F_X^{A,Y}$ respectively, then for $1 \leq i \leq M$, $s_i = \bar{s_i}$.
First, by the definition of invariants factors, for any choice of X, Y

$$dim(span(X, A^{tr}X, (A^{tr})^2X, (A^{tr})^3X, \cdots)) \leq \sum_{i=1}^{M} deg(s_i)$$
$$dim(span(Y, AY, A^2Y, A^3Y, \ldots)) \leq \sum_{i=1}^{M} deg(s_i)$$

(Note: equality is possible for both inequalities). Then define two block Hankel matrices (with as many rows and columns as it takes to maximize the rank)

$$H_Y = \begin{bmatrix} I \\ A \\ A^2 \\ A^3 \\ \vdots \end{bmatrix} \begin{bmatrix} Y & AY & A^2Y & \cdots \end{bmatrix}$$

$$H_{X,Y} = \begin{bmatrix} X^{tr} \\ X^{tr}A \\ X^{tr}A^2 \\ X^{tr}A^3 \\ \vdots \end{bmatrix} \begin{bmatrix} Y & AY & A^2Y & A^3Y & \cdots \end{bmatrix}$$

Now,

$$rank(H_Y) = dim(span(Y, AY, A^2Y, A^3Y, \dots))$$

and

$$rank(H_{X,Y}) = dim(span(X^{tr}Y, X^{tr}AY, X^{tr}A^2Y, \cdots)) \leq rank(H_Y)$$

Let W, Z be generic choices for X, Y respectively. A generic choice of Y maximizes $dim(span(Y, AY, A^2Y, A^3Y, \dots))$; thus, we get

$$rank(H_Z) = dim(span(Z, AZ, A^2Z, A^3Z, \dots)) = \sum_{i=1}^{M} deg(s_i)$$

A generic choice of X makes $F_X^{A,Y} = F^{A,Y}$ and by (2.6), $rank(H_{W,Z}) = deg(det(F_W^{A,Z})) = deg(det(F^{A,Z})) = rank(H_Z)$. Therefore,

$$deg(det(F_W^{A,Z})) = \sum_{i=1}^{M} deg(s_i)$$

Since $\bar{s}_i$ divides s_i, in order to have $deg(det(F_W^{A,Z})) = \sum_{i=1}^{M} deg(s_i)$, it must be the case that $\bar{s}_i = s_i$ for $1 \leq i \leq M$ as needed.